

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 1 437 409 A1

(12)

EUROPEAN PATENT APPLICATION

published in accordance with Art. 158(3) EPC

(43) Date of publication:

14.07.2004 Bulletin 2004/29

(51) Int Cl.7: **C12N 15/29, C12Q 1/68**

(21) Application number: **02799470.6**

(86) International application number:

PCT/JP2002/009429

(22) Date of filing: **13.09.2002**

(87) International publication number:

WO 2003/027290 (03.04.2003 Gazette 2003/14)

(84) Designated Contracting States:

**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR
IE IT LI LU MC NL PT SE SK TR**

Designated Extension States:

AL LT LV MK RO SI

(72) Inventors:

- **KOMORI, Toshiyuki, Japan Tobacco Inc.
Iwata-gun, Shizuoka 438-0802 (JP)**
- **OTA, Shozo, Japan Tobacco Inc.
Yokohama-shi, Kanagawa 227-8512 (JP)**
- **MURAI, Nobuhiko, Japan Tobacco Inc.
Tokyo 105-8422 (JP)**
- **HIEI, Yuko, Japan Tobacco Inc.
Iwata-gun, Shizuoka 438-0802 (JP)**

(30) Priority: **19.09.2001 JP 2001285247**

04.10.2001 JP 2001309135

26.06.2002 JP 2002185709

(74) Representative: **Behnisch, Werner, Dr.**

Reinhard-Skühra-Weise & Partner,

Patentanwälte,

Postfach 44 01 51

80750 München (DE)

(71) Applicants:

- **Japan Tobacco Inc.
Tokyo 105-8422 (JP)**
- **Syngenta Limited
Guildford Surrey GU2 7YH (GB)**

(54) **METHOD OF IMPARTING OR CONTROLLING FERTILITY WITH THE USE OF FERTILITY RESTORING GENE FOR RICE BT-MALE STERILITY CYTOPLASM AND METHOD OF JUDGING THE EXISTENCE OF FERTILITY RESTORING GENE**

(57) The purpose of the present invention is to provide a method for providing and inhibiting the rice fertility to the rice BT type cytoplasmic male sterility, and discerning the presence of the rice restorer gene. The present invention employs a nucleic acid having the base sequence of SEQ ID NO.27, or a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27, and which func-

tions to restore fertility. Alternatively, a nucleic acid having the base sequence of bases 38538-54123 of SEQ ID NO.27, or having a base sequence which is identical to at least 70% of the base sequence of bases 38538-54123 of SEQ ID NO.27, and which functions to restore fertility, is used.

EP 1 437 409 A1

Description

FIELD OF THE INVENTION

5 [0001] The present invention relates to a method for providing and inhibiting the rice fertility, and discerning the presence of the rice restorer gene by using the rice BT type cytoplasmic male sterility.

[0002] The present application claims priority based on Japanese Patent Application No. 2001-285247 filed on September 19, 2001, Japanese Patent Application No. 2001-309135 filed on October 4, 2001 and Japanese Patent Application No. 2002-185709 filed on June 26, 2002. The entire disclosures of the three patent applications are incorporated herein.

PRIOR ART

15 [0003] Rice is a self-fertilizing plant, so in order to perform crossing between varieties, self-fertilization must first be avoided by removing all stamens in a glumaceous flower just before flowering and, then fertilization is effected with pollens from the parent variety with which it is to be crossed. However, this manual crossing method is entirely unsuitable for producing a large quantity of hybrid seeds on a commercial scale.

[0004] Accordingly, hybrid rice is produced by the triple-crossing system which makes use of cytoplasmic male sterility. In the triple-crossing system, the following three lines are employed, i.e., a sterile line having male sterile cytoplasm, a restorer line having Rf-1 gene and a maintainer line having the same nuclear gene as that of the sterile line but not having any sterile cytoplasm. By using these three lines, (i) hybrid seeds can be obtained through fertilization of the sterile line with the pollen of the restorer line whereas (ii) the sterile line can be maintained through its fertilization with the pollen of the maintainer line.

20 [0005] When employing the BT type male sterile cytoplasm in the triple-crossing system, it is important to breed rice of the restorer line and to this end, it is necessary to ensure that the rice at every stage of breeding maintains Rf-1 gene and that the Rf-1 gene is homozygous at the final stage. It also becomes necessary in the triple-crossing system to check to ensure that the variety used as the restorer line possesses Rf-1 gene, or to check for the presence of Rf-1 gene in order to ensure that the resulting hybrid seeds have restored fertility.

25 [0006] In order to genotype the locus of Rf-1 gene in a plant, it has been necessary that F1 plants be first formed from hybrid seeds obtained by crossing the plant to be genotyped to a standard line and then self-fertilized, followed by investigating the incidence of individuals that can produce seeds at a frequency higher than a certain level (e.g. 70~80% or more). The standard line refers to the maintainer line, the sterile line or a set of the two lines, and it is appropriately chosen depending upon whether the cytoplasm of the individual under test is of BT type or normal type or unknown. If the standard line is a sterile line, it is crossed to the individual under test as the female parent and if the standard line is a maintainer line, it is crossed as the male parent.

30 [0007] However, these techniques require a huge amount of labor and time to carry out. As a further problem, fertilization for seed production is sensitive to environmental factors and if an investigation is made in an unfavorable environment such as cold climate or insufficient daylight, sterility may be caused irrespective of the genotype constitution, with the result that genotyping of the locus of Rf-1 gene cannot be performed accurately.

35 [0008] With a view to solving these problems, it has recently been proposed that Rf-1 gene be checked for its presence by a technique of molecular biology. The technical idea of this technique lies in checking for the presence or absence of Rf-1 gene by detecting base sequences linked to Rf-1 gene (such sequences are hereunder referred to as DNA markers). Note that it is not possible to directly detect Rf-1 gene since the DNA sequence of Rf-1 gene has not been clarified so far.

40 [0009] For example, it has been reported that the locus of Rf-1 gene in rice is present on chromosome 10 and located between DNA marker (RFLP marker) loci G291 and G127 which can be used in restriction fragment length polymorphism analysis (RFLP) (Fukuta et al., 1992, Jpn J. Breed. 42 (supl. 1) 164-165). This is a known method of genotyping the locus of Rf-1 gene by investigating the genotypes of DNA marker loci G291 and G127 which are linked to Rf-1 gene.

45 [0010] However, the conventional molecular biology techniques have several problems. First, they use RFLP markers which need to be detected by Southern blot analysis. In order to perform Southern blot analysis, DNA at the microgram level needs to be purified from the individual under test and, in addition, there is a need to carry out a sequence of steps comprising treatment with restriction enzymes, electrophoresis, blotting, hybridization with a probe and signal detection; this not only involves considerable labor but it also takes about one week to obtain the test results.

50 [0011] The second problem is that since the gene map distance between RFLP marker loci G291 and G127 is as long as about 30 cM (corresponding to about 9000 kbp in rice DNA), the probability for the occurrence of double recombination in the region would be a few percent and hence, it is not always guaranteed that the genotype of the locus of Rf-1 gene can be estimated correctly by the markers.

55 [0012] Thirdly, when the presence of Rf-1 gene is estimated by detecting RFLP marker loci G291 and G127, not only

Rf-1 gene but also the gene region between those loci are introduced into the fertility restorer line selected as the result of breeding. As a consequence, the introduced DNA sequence will have a chromosomal region of 30 cM or longer from the Rf-1 gene donor parent, and this presents the risk of introducing a deleterious gene that may potentially be present within that region.

[0013] In order to solve these problems, there have been developed a dominant DNA marker (Japanese Patent Public Disclosure No. 222588/1995) and a co-dominant DNA marker (Japanese Patent Public Disclosure No. 313187/1997), both of which are linked to the locus of Rf-1 gene. These markers are linked to the locus of Rf-1 gene, their genetic distances from Rf-1 gene respectively being 1.6 ± 0.7 cM (corresponding to about 480 kbp in rice DNA) and 3.7 ± 1.1 cM (corresponding to about 1110 kbp in rice DNA), and their loci being on opposite sides of the locus of Rf-1 gene. Hence, the presence of Rf-1 gene can be estimated by detecting the presence of both the locus of the dominant PCR marker and that of the co-dominant PCR marker. The use of the co-dominant PCR marker also enables us to estimate as to whether the locus of Rf-1 gene is homozygous or heterozygous.

[0014] However, the use of these PCR markers still involve several problems. The co-dominant marker has a genetic distance of 3.7 ± 1.1 cM from the locus of Rf-1 gene, and the problem of potentially high frequency of recombination with the locus of Rf-1 gene has not been fully dissolved. As a result, speaking of the co-dominant marker itself, correct detection can be made as to whether it is homozygous or a heterozygous. However, if recombination occurs between the locus of the co-dominant marker and that of Rf-1 gene, the genotype of Rf-1 gene locus cannot be determined correctly, particularly as to whether it is homozygous or heterozygous. On the other hand, if the dominant marker is used to genotype the locus of Rf-1 gene, the marker will detect individuals indiscriminately irrespective of whether they are homozygous (Rf-1/Rf-1) or heterozygous (Rf-1/rf-1) with respect to Rf-1 gene. Therefore, even if the co-dominant marker is used in combination with the dominant marker in order to genotype the locus of Rf-1 gene, it is not possible to correctly distinguish individuals having Rf-1 gene homozygously from those having the gene heterozygously. Further, if no amplification product is obtained in PCR using the dominant marker, one cannot deny the possibility that this is due to some problems in the experimental procedure. As a further problem, since the genetic distance between the co-dominant marker and the dominant marker is as great as about 5.3 cM (around 1590 kbp), the size of the chromosomal region introduced from the Rf-1 gene donor parent cannot be limited to a sufficiently small value to prevent any concomitant introduction of a deleterious gene which may be contained in that region.

[0015] Japanese Patent Public Disclosure No. 139465/2000 describes co-dominant PCR markers that were developed on the basis of the base sequences of RFLP markers located in the neighborhood of Rf-1 gene on chromosome 10 of rice. However, most of those PCR markers are spaced from the Rf-1 gene by a genetic distance greater than about 1 cM.

SUMMARY OF THE INVENTION

[0016] An object of the present invention is to provide methods for restoring rice fertility. A method of the present invention comprises introducing a nucleic acid into rice, wherein the nucleic acid has the base sequence of SEQ ID NO.27, or has a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27, and which functions to restore fertility. Another method of the present invention comprises introducing a nucleic acid into rice, wherein the nucleic acid has the base sequence of bases 38538-54123 of SEQ ID NO.27, or has a base sequence which is identical to at least 70% of the base sequence of bases 38538-54123 of SEQ ID NO.27, and which functions to restore fertility. Still another method of the present invention comprises introducing a nucleic acid into rice, wherein the nucleic acid has the base sequence of bases 42357-53743, more preferably, bases 42132-48883 of SEQ ID NO.27, or has a base sequence which is identical to at least 70% of the base sequence of bases 42357-53743, more preferably, bases 42132-48883 of SEQ ID NO.27, and which functions to restore fertility. In an embodiment of the methods of the present invention, a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27 or of the base sequence of bases 38538-54123 of SEQ ID NO.27 meets at least one of the following requirements 1) and 2):

- 1) a base corresponding to the base 45461 of SEQ ID NO.27 is A; and
- 2) a base corresponding to the base 49609 of SEQ ID NO.27 is A.

[0017] Another object of the present invention is to provide a method for discerning whether or not a subject rice individual or a seed thereof has the Rf-1 gene or not. The discerning method of the present invention utilizes a fact that a sequence determining the presence of the function of the rice restorer gene (the Rf-1 gene) positions between the polymorphism detection marker loci P4497 Mbol and B56691 Xab I on rice chromosome 10.

[0018] In an embodiment of the methods of the present invention, the subject rice individual or the seed thereof is determined to have the Rf-1 gene, in the case that the nucleic acid having a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27 or of the base sequence of bases 38538-54123 of SEQ ID NO.27,

meets at least one of the following requirements 1) and 2):

- 1) a base corresponding to the base 45461 of SEQ ID No.27 is A; and
- 2) a base corresponding to the base 49609 of SEQ ID NO.27 is A.

[0019] Another object of the present invention is to provide a method for inhibiting the function of the Rf-1 gene to restore fertility. The inhibition method of the present invention comprises, in an embodiment, introducing an antisense having at least 100 continuous bases in length, and having a base sequence complementary to a nucleic acid having the base sequence of SEQ ID NO.27, or to a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27, and which functions to restore fertility. In another embodiment, the inhibition methods of the present invention comprise introducing an antisense having at least 100 continuous bases in length, and having a base sequence complementary to a nucleic acid having the base sequence of bases 38538-54123 of SEQ ID NO.27, or to a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of bases 38538-54123 of SEQ ID NO.27, and which functions to restore fertility.

[0020] Another object of the present invention is to provide a nucleic acid having the base sequence of SEQ ID NO. 27, or a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO. 27, and which functions to restore fertility. The present invention also provides a nucleic acid having the base sequence of bases 38538-54123 of SEQ ID NO.27, or a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of bases 38538-54123 of SEQ ID NO.27, and which functions to restore fertility. The present invention also provides a nucleic acid having the base sequence of bases 42357-53743, more preferably, bases 42132-48883 of SEQ ID NO.27, or a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of bases 42357-53743, more preferably, bases 42132-48883 of SEQ ID NO.27, and which functions to restore fertility.

BRIEF DESCRIPTION OF THE DRAWING

[0021]

Fig. 1 shows the results of chromosomal walking started from the RFLP marker locus S12564.

Fig. 2 shows an alignment of lambda clone contigs in relation to the BAC clone AC068923.

Fig. 3 shows the chromosomal organization of recombinant pollens proximal to the Rf-1 locus (all fertile) as mapped in close proximity to the Rf-1 locus based on the genotypes at the marker loci of 10 individuals (RS1, RS2, RC1-8) generated from the pollens. White bars represent japonica regions and black bars represent indica regions.

Fig. 4 is a gene map in which the locus of Rf-1 gene on chromosome 10 of rice is positioned on a linkage map in relation to various markers; the values of map distance were calculated from the segregation data from 1042 F1 individuals.

Fig. 5 shows fragments from 10 genomic clones used for the identification of the Rf-1 region by complementation assays. Lambda clones obtained by chromosomal walking (thin lines) were used for complementation assays of the chromosomal regions shown by bold lines. XSF18 was found to contain a deletion shown by dotted line.

Fig. 6 shows the results of complementation assays using a 15.7 kb fragment from XSG16 (Example 10) and a 16.2 kb fragment from XSF18 (Example 8). The plant transformed with the 15.7 kb fragment from XSG16 has restored fertility as proved by ears bowing.

BEST MODES FOR PERFORMING THE INVENTION

[0022] We began by restricting the Rf-1 locus to a very small region on chromosome 10. On this basis, we developed PCR markers proximal to the Rf-1 locus and found a method for detecting the Rf-1 gene by utilizing on the linkage of these PCR markers to the Rf-1 locus. Specifically, the presence of the Rf-1 gene is tested and individuals homozygous for the Rf-1 gene are selected by genotyping at the novel PCR marker loci proximal to the Rf-1 locus on the basis that the Rf-1 locus is mapped between the PCR marker loci S12564 Tsp509I and C1361 MwoI on chromosome 10 of rice. We previously filed a patent application for the method for detecting the Rf-1 gene under Japanese Patent Application No. 2000-247204 on August 17, 2000. The entire disclosure of the patent application is incorporated herein by reference.

I. Methods for estimating the genotype at the Rf-1 locus described in Japanese Patent Application No. 2000-247204

[0023] Japanese Patent Application No. 2000-247204 describes methods for determining whether or not a rice individual or seed under test has the Rf-1 gene on the basis that the Rf-1 locus is mapped between the PCR marker loci

S12564 Tsp509I and C1361 MwoI on chromosome 10 of rice.

Markers

[0024] Primer pairs designed to be specific to particular regions near the locus of Rf-1 gene are used in PCR and the amplification products are treated with particular restriction enzymes; upon electrophoresis, rice of indica lines in some cases provide an observable band of a different size from that of rice of Japonica lines. This band which is characteristic of indica lines is herein referred to as the Rf-1 linked band. Now that it has been made clear by the present inventors that the locus of Rf-1 gene is located between PCR markers S12564 Tsp509I and C1361 MwoI on chromosome 10 of rice, the skilled artisan can appropriately develop and employ PCR markers that are present in the neighborhood of Rf-1 gene.

[0025] For instance, according to the invention, a rice individual under test is checked to see if its genome contains at least one of the PCR markers listed below, thereby determining whether the individual under test has Rf-1 gene linked to those PCR markers:

(1) marker 1: PCR marker R1877 EcoRI which, when rice genomic DNA is subjected to PCR with DNA primers having the sequences of SEQ ID NO:1 and SEQ ID NO:2, can detect polymorphisms between rice individuals of the japonica and indica lines depending on whether the amplification products have a recognition site for restriction enzyme EcoRI;

(2) marker 2: PCR marker G4003 HindIII (SEQ ID NO:19) which, when rice genomic DNA is subjected to PCR with DNA primers having the sequences of SEQ ID NO:3 and SEQ ID NO:4, can detect polymorphisms between rice individuals of the japonica and indica lines depending on whether the amplification products have a recognition site for restriction enzyme HindIII;

(3) marker 3: PCR marker C1361 MwoI (SEQ ID NO:20) which, when rice genomic DNA is subjected to PCR employing DNA primers having the sequences of SEQ ID NO:5 and SEQ ID NO:6, can detect polymorphisms between rice individuals of the japonica and indica lines depending on whether the amplification products have a recognition site for restriction enzyme MwoI;

(4) marker 4: PCR marker G2155 MwoI (SEQ ID NO:21) which, when rice genomic DNA is subjected to PCR with DNA primers having the sequences of SEQ ID NO:7 and SEQ ID NO:8, can detect polymorphisms between rice individuals of the japonica and indica lines depending on whether the amplification products have a recognition site for restriction enzyme MwoI;

(5) marker 5: PCR marker G291 MspI (SEQ ID NO:22) which, when rice genomic DNA is subjected to PCR with DNA primers having the sequences of SEQ ID NO:9 and SEQ ID NO:10, can detect polymorphisms between rice individuals of the japonica and indica lines depending on whether the amplification products have a recognition site for restriction enzyme MspI;

(6) marker 6: PCR marker R2303 BstI (SEQ ID NO:23) which, when rice genomic DNA is subjected to PCR with DNA primers having the sequences of SEQ ID NO:11 and SEQ ID NO:12, can detect polymorphisms between rice individuals of the japonica and indica lines depending on whether the amplification products have a recognition site for restriction enzyme BstI;

(7) marker 7: PCR marker S10019 BstUI (SEQ ID NO:24) which, when rice genomic DNA is subjected to PCR with DNA primers having the sequences of SEQ ID NO:13 and SEQ ID NO:14, can detect polymorphisms between rice individuals of the japonica and indica lines depending on whether the amplification products have a recognition site for restriction enzyme BstUI;

(8) marker 8: PCR marker S10602 KpnI (SEQ ID NO:25) which, when rice genomic DNA is subjected to PCR with DNA primers having the sequences of SEQ ID NO:15 and SEQ ID NO:16, can detect polymorphisms between rice individuals of the japonica and indica lines depending on whether the amplification products have a recognition site for restriction enzyme KpnI; and

(9) marker 9: PCR marker S12564 Tsp509I (SEQ ID NO:26) which, when rice genomic DNA is subjected to PCR with DNA primers having the sequences of SEQ ID NO:17 and SEQ ID NO:18, can detect polymorphisms between rice individuals of the japonica and indica lines depending on whether the amplification products have a recognition site for restriction enzyme Tsp509I.

[0026] Assuming that the locus of Rf-1 gene was highly likely to be located near the nine RFLP marker regions R1877, G291, R2303, S12564, C1361, S10019, G4003, S10602 and G2155 on chromosome 10 of rice (see the results of RFLP linkage analysis described in Fukuta et al., 1992, Jpn. J. Breed. 42 (supl. 1) 164-165 and the RFLP linkage map of rice described in Harushima et al., 1998, Genetics, 148, 479-494), the present inventors converted those RFLP markers to co-dominant PCR markers such as CAPS markers or dCAPS markers as described below in Reference example 1 (Michaels and Amasino, 1998, The Plant Journal, 14(3), 381-385; Neff et al., 1998, The Plant Journal, 14

(3), 387-392). As a result of this conversion, the PCR markers above have been obtained.

[0027] Among these PCR markers, one group consisting of PCR markers R1877 EcoRI, G291 MspI (SEQ ID NO: 22), R2303 BstI (SEQ ID NO:23) and S12564 Tsp509I (SEQ ID NO:26) and the other group consisting of PCR markers C1361 MwoI (SEQ ID NO:20), S10019 BstUI (SEQ ID NO:24), G4003 HindIII (SEQ ID NO:19), S10602 KpnI (SEQ ID NO:25) and G2155 MwoI (SEQ ID NO:21) are on opposite sides of the locus of Rf-1 gene on chromosome 10 of rice.

[0028] Therefore, in one embodiment, the presence of the Rf-1 gene is detected by detecting Rf-1 linked bands by (a) at least one PCR marker selected from the group consisting of PCR markers R1877 EcoRI, G291 MspI, R2303 BstI and S12564 Tsp509I, and (b) at least one PCR marker selected from the group consisting of PCR markers C1361 MwoI, S10019 BstUI, G4003 HindIII, S10602 KpnI and G2155 MwoI. In this case, at least S12564 Tsp509I from group (a) and at least C1361 MwoI from group (b) are preferably used as the closest PCR markers to the Rf-1 gene. If Rf-1 linked bands are detected with PCR markers of both (a) and (b) in the genome of the rice under test, it can be estimated with a high probability that the rice contains Rf-1 gene.

[0029] In another embodiment, Rf-1 linked bands are detected by at least two PCR markers of group (a) and at least two PCR markers of group (b) above. For example, a rice individual carrying the Rf-1 gene with a minimum of unwanted gene regions can be selected by picking up an individual in which Rf-1 linked bands are detected by markers of groups (a) and (b) more proximal to the Rf-1 gene but not detected by markers of groups (a) and (b) more distal from the Rf-1 gene on the gene map shown in Fig. 1. Again, it is preferred that at least one PCR marker of group (a) is S12564 Tsp509I and at least one PCR marker of group (b) is C1361 MwoI. Thus, the two PCR marker loci S12564 Tsp509I and C1361 MwoI are separated by a genetic distance of 0.3 cM. By utilizing this characteristic, the chromosomal region that is introduced from the Rf-1 gene donor parent can be narrowed down to a size of about 1 cM. This helps minimize the possibility of introducing into the restorer line a deleterious gene that may be present in the neighborhood of Rf-1 gene in the donor parent.

Detection of the Rf-1 gene

[0030] In order to detect Rf-1 gene in the genome of a rice under test, any one of the above PCR markers is amplified from the genome of the rice by PCR using primers of SEQ ID NOS: 1-18 above and then detected by the polymerase chain reaction-restriction fragment length polymorphism method (PCR-RFLP). PCR-RFLP is a method that is applicable to the case where polymorphisms exist among variety lines at recognition sites of restriction enzymes in the sequences of PCR amplified DNA fragments and by which specific polymorphisms can conveniently be identified on the basis of cleavage patterns with those restriction enzymes (D.E. Harry et al., Theor. Appl. Genet. (1998), 97:327-336)

[0031] Restriction enzyme cleavage patterns show the bands as shown in Table 1 below on a visualized gel depending on the primer pair used.

Table 1

	Approximate size (bp) of detected band
Detection of marker 1 (R1877 EcoRI) with primer pair 1	
When the genome of test rice has Rf-1 gene homozygously	1500 and 1700
When the genome of test rice has Rf-1 gene heterozygously	1500, 1700 and 3200
When the genome of test rice has no Rf-1 gene	3200
Detection of marker 2 (G4003 HindIII) with primer pair 2	
When the genome of test rice has Rf-1 gene homozygously	362
When the genome of test rice has Rf-1 gene heterozygously	95, 267 and 362
When the genome of test rice has no Rf-1 gene	95 and 267
Detection of marker 3 (C1361 MwoI) with primer pair 3	
When the genome of test rice has Rf-1 gene homozygously	50 and 107
When the genome of test rice has Rf-1 gene heterozygously	25, 50, 79 and 107
When the genome of test rice has no Rf-1 gene	25, 50 and 79
Detection of marker 4 (G2155 MwoI) with primer pair 4	
When the genome of test rice has Rf-1 gene homozygously	25, 27 and 78
When the genome of test rice has Rf-1 gene heterozygously	25, 27, 78 and 105
When the genome of test rice has no Rf-1 gene	25 and 105

Table 1 (continued)

	Approximate size (bp) of detected band
Detection of marker 5 (G291 MspI) with primer pair 5	
When the genome of test rice has Rf-1 gene homozygously	25, 49 and 55
When the genome of test rice has Rf-1 gene heterozygously	25, 49, 55 and 104
When the genome of test rice has no Rf-1 gene	25 and 104
Detection of marker 6 (R2303 BstI) with primer pair 6	
When the genome of test rice has Rf-1 gene homozygously	238, 655 and 679
When the genome of test rice has Rf-1 gene heterozygously	238, 655, 679 and 1334
When the genome of test rice has no Rf-1 gene	238 and 1334
Detection of marker 7 (S10019 BstUI) with primer pair 7	
When the genome of test rice has Rf-1 gene homozygously	130, 218 and 244
When the genome of test rice has Rf-1 gene heterozygously	130, 218, 244 and 462
When the genome of test rice has no Rf-1 gene	130 and 462
Detection of marker 8 (S10602 KpnI) with primer pair 8	
When the genome of test rice has Rf-1 gene homozygously	724
When the genome of test rice has Rf-1 gene heterozygously	117, 607 and 724
When the genome of test rice has no Rf-1 gene	117 and 607
Detection of marker 9 (S12564 Tsp509I) with primer pair 9	
When the genome of test rice has Rf-1 gene homozygously	41 and 117
When the genome of test rice has Rf-1 gene heterozygously	26, 41, 91 and 117
When the genome of test rice has no Rf-1 gene	26, 41 and 91

II. Identification of the Rf-1 locus

[0032] As described above, Japanese Patent Application No. 2000-247204 discloses RFLP-PCR markers based on our finding that the Rf-1 locus is mapped between DNA marker loci S12564 Tsp509I and C1361 MwoI. Fertility-restoring lines are established by backcrossing the Rf-1 gene into a normal japonica variety not containing the Rf-1 gene. If the method for identifying the Rf-1 locus described in Japanese Patent Application No. 2000-247204 is used during this process, not only the restoring lines can be established efficiently (within 2-3 years) but also the length of insert fragments can be controlled.

[0033] However, introduction by crossing inevitably introduce regions proximal to Rf-1 at the same time. Japanese Patent Application No. 2000-247204 showed that the Rf-1 locus is mapped between DNA marker loci S12564 Tsp509I and C1361 MwoI, but the distance between both loci is about 0.3 cM, i.e. about 90 kbp. If a deleterious gene existed proximal to Rf-1, it would be undeniable that the deleterious gene might be inserted together with the Rf-1 gene.

[0034] Thus, we searched for regions linked to the Rf-1 gene between DNA marker loci S12564 Tsp509I and C1361 MwoI by chromosomal walking and genetic analysis based on the close linkage between the Rf-1 locus and the DNA marker locus S12564 Tsp509I. As a result, we successfully identified the region of the Rf-1 locus including the Rf-1 gene upto about 76 kb and determined the entire base sequence of said region. According to the present invention, it is possible to introduce the function of a fertility restorer gene into BT male sterile cytoplasms by genetic engineering techniques.

[0035] Specifically, in Japanese Patent Application No. 2000-247204, linkage analyses on a population of 1042 individuals prepared by pollinating MS Koshihikari with MS-FR Koshihikari (heterozygous at the Rf-1 locus) revealed one recombinant between the Rf-1 and S12564 Tsp509I loci and two recombinants between the Rf-1 and C1361 MwoI loci (Reference examples 1-2 herein). In the present invention, 4103 individuals were added to the population to analyze a total of 5145 individuals. As a result, one recombinant between the Rf-1 and S12564 Tsp509I loci and six recombinants between the Rf-1 and C1361 MwoI loci were newly found with a total of 2 and 8 recombinants. These 10 individuals were tested by the high-precision segregation analysis of the present invention as recombinants proximal to the Rf-1 locus (Example 1).

[0036] The frequency of 8 recombinants between the Rf-1 and C1361 MwoI loci as compared with 2 recombinants between the Rf-1 and S12564 Tsp509I loci means that the S12564 Tsp509I locus is genetically closer to the Rf-1 locus

than the C1361 Mwo1 locus. Genetic distance (expressed in recombination frequency: cM) and physical distance (expressed in the number of base pairs: bp) are not always proportional to each other, but it can be normally expected that physical distance decreases with genetic distance.

[0037] Thus, we tried to isolate the Rf-1 locus by chromosomal walking started from the S12564 Tsp509I locus (Example 2). Chromosomal walking was performed on a genomic library prepared from λ DASH II vector using the genomic DNA of an indica variety IR24 and a japonica variety Asominori. IR24 is a variety carrying Rf-1, while Asominori is a variety not carrying Rf-1. As a result of chromosomal walking, contigs covering a chromosomal region of about 76 kb (ordered sets of overlapping clones on a chromosome) were able to be prepared from genomic clones of IR24, and the entire base sequence (76363 bp) thereof was determined.

[0038] Then, 12 markers were newly developed on the basis of the base sequence data or the like obtained and a high-precision segregation analysis was performed on the 10 recombinants proximal to Rf-1 locus described above (Example 3). As a result, a 65 kb sequence included in the chromosomal region of about 76 kb above was shown to contain a sequence determining the presence of the function of the Rf-1 gene. This region is covered by a contig consisting of 8 genomic clones. Each clone has a length of about 12-22 kb and has overlapping domains of at least 4.7 kb. Genes for rice are known to have a wide range of lengths (from short ones to large ones), but most of them seem to have a length of several kbs or less. Thus, at least one of these 8 genomic clones is expected to contain the full-length Rf-1 gene.

[0039] We further restricted the Rf-1 gene region in the chromosomal region of about 76 kb above and performed complementation assays to directly demonstrate the presence of a fertility restoring ability.

[0040] Specifically, 10 partial fragments (each 10-21 kb) in the above region of 76 kb were separately introduced into immature seeds of a male sterility line MS Koshihikari by genetic engineering techniques (Fig. 5). Of the 10 partial fragments used, 8 fragments are derived from 8 genomic clones previously obtained by chromosomal walking (XSE1, XSE7, XSF4, XSF20, XSG22, XSG16, XSG8 and XSH18 shown in Fig. 1 and described in Example 3). Additionally, fragments derived from 2 clones XSF18 and XSX1 were also analyzed by complementation assays. XSF18 is identical to XSF20 at the 5' and 3' ends (bases 20328 and 41921 of SEQ ID NO: 27, respectively), but lacks internal bases 33947-38591. This is because clone XSF18 was initially isolated but found to contain the above deletion during amplification after isolation, and therefore, the amplification step was freshly taken to isolate a complete clone designated XSF20 (Example 8). XSX1 is a clone freshly prepared from clones XSG8 and XSH18 by restriction enzyme treatment and ligation to contain sufficient overlapping domains because of the overlapping domains of both clones are relatively small (about 7 kb) (Example 13).

[0041] If the insert fragment completely contains the Rf-1 gene, transformed individuals at this generation restore fertility because Rf-1 is a dominant gene. In complementation assays plants transformed with each fragment were evaluated for seed fertility to find that those transformed with a 15.6 kb fragment (including bases 38538-54123 of SEQ ID NO: 27) derived from the λ phage clone XSG16 restored seed fertility (Example 10). Plants transformed with the other fragments were all sterile. These results showed that the above 15.6 kb fragment completely contains the Rf-1 gene. Moreover, a method for introducing the Rf-1 gene by genetic engineering techniques was provided by the present invention and demonstrated to be effective.

[0042] To further specify the region of the λ phage clone XSG16 in which the Rf-1 gene is contained, we evaluated seed fertility of shorter fragments than the 15.6 kb fragment (including bases 38538-54123 of SEQ ID NO: 27) by complementation assays. As a result, plants transformed with a 11.4 kb fragment derived from XSG16 (including bases 42357-53743 of SEQ ID NO: 27) were shown to restore seed fertility (Example 10(2)). Plants transformed with a further shorter 6.8 kb fragment (including bases 42132-48883 of SEQ ID NO: 27) also restored seed fertility (Example 10(3)). These results showed that the above 6.8 kb fragment contains the Rf-1 gene.

III. Nucleic acids containing the Rf-1 locus

[0043] The present invention provides nucleic acids containing the locus of a fertility restorer gene (Rf-1). The nucleic acids containing the locus of a fertility restorer gene (Rf-1) of the present invention include a nucleic acid having the base sequence of SEQ ID NO.27, or a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27, and which functions to restore fertility.

[0044] As described in Example 10, it was confirmed that the Rf-1 gene is completely contained in especially bases 38538-54123 of the base sequence of SEQ ID NO: 27. Accordingly, the present invention especially provides a nucleic acid having the base sequence of bases 38538-54123 of SEQ ID NO.27, or has a base sequence which is identical to at least 70% of the base sequence of bases 38538-54123 of SEQ ID NO.27, and which functions to restore fertility.

As used herein, the term "the base sequence of SEQ ID NO: 27" refers to the entire sequence of SEQ ID NO: 27 or a part thereof participating in fertility restoring function, especially bases 38538-54123 according to the context. More preferably, it refers to bases 42357-53743, still more preferably bases 42132-48883.

[0045] In the examples below, a nucleic acid was isolated from a genomic library of indica rice IR24 containing the

Rf-1 gene as a nucleic acid containing a fertility restorer gene (Rf-1) and determined to have the base sequence of SEQ ID NO: 27. However, the nucleic acid containing a fertility restorer gene (Rf-1) of the present invention can be derived from any indica variety carrying the RF-1 gene. The indica varieties carrying the Rf-1 gene include, but not specifically limited to, e.g. IR24, IR8, IR36, IR64, Chinsurah and Boroli. Known japonica varieties not carrying the Rf-1 gene include, but not limited to, Asominori, Koshihikari, Kirara 397, Akihikari, Akitakomachi, Sasanishiki, Kinuhikari, Nipponbare, Hatsuboshi, Koganebare, Hinohikari, Mineasahi, Aichinokaori, Hatsushimo, Akebono, Fujihikari, Minenoyukimochi, Kokonoemochi, Fukuhibiki, Dontokoi, Gohyakumangoku, Hanaechizen, Todorokiwase, Haenuki, Domannaka, Yamakikari, etc. The "indica" and "japonica" varieties are well known to those skilled in the art and the rice varieties encompassed by the present invention can be readily determined by those skilled in the art.

[0046] Nucleic acids of the present invention include DNA in both single-stranded and double-stranded forms, as well as the RNA complement thereof. DNA includes, for example, genomic DNA (including corresponding cDNA), chemically synthesized DNA, DNA amplified by PCR, and combinations thereof.

[0047] Nucleic acids containing the Rf-1 gene of the present invention preferably have the base sequence of SEQ ID NO: 27. More than one codon may encode the same amino acid, and this is called degeneracy of the genetic code. Thus, a DNA sequence not completely identical to SEQ ID NO: 27 may encode a protein having an amino acid sequence completely identical to SEQ ID NO: 27. Such a variant DNA sequence may result from silent mutation (e.g., occurring during PCR amplification), or can be a product of deliberate mutagenesis of a native sequence.

[0048] It is well known for those skilled in the art that even proteins having the same function may have different amino acid sequences depending on the varieties from which they are derived. The Rf-1 gene of the present invention includes such homologs and variants of the base sequence of SEQ ID NO: 27 so far as they functions to restore fertility. The expression "function to restore fertility" means that fertility is conferred on a rice individual or seed when such a DNA fragment is introduced. Fertility restoration may result from the expression of a protein by the Rf-1 gene or some function of the nucleic acid (DNA or RNA) per se of the Rf-1 gene in conferring fertility.

[0049] Whether or not a homolog or variant of the Rf-1 gene functions to restore fertility can be examined by, but not limited to, the following method, for example. A nucleic acid fragment under test is introduced into immature seeds obtained by pollinating MS Koshihikari (sterile line) with MS-FR Koshihikari according to the method of Hiei et al. (Plant Journal (1994), 6(2), p. 272-282). As the resulting transformants are cultured under normal conditions, the seeds mature only when the nucleic acid fragment under test functions to restore fertility.

[0050] The nucleic acid derived from a corresponding region of japonica Asominori not carrying the Rf-1 gene has the base sequence shown in SEQ ID NO: 28. Corresponding parts of SEQ ID NO: 28 and SEQ ID NO: 27 have an overall identity of about 98%. Thus, nucleic acids containing the locus of a fertility restorer gene (Rf-1) of the present invention are at least about 70%, preferably about 80% or more, more preferably 90% or more, still more preferably 95% or more, most preferably 98 or more% identical to SEQ ID NO: 27.

[0051] The percent identity may be determined by visual inspection and mathematical calculation. Alternatively, the percent identity of two nucleic acid sequences can be determined by comparing sequence information using the GAP computer program, version 6.0 described by Devereux et al., Nucl. Acids Res., 12:387 (1984) and available from the University of Wisconsin Genetics Computer Group (UWGCG). The preferred default parameters for the GAP program include: (1) a unary comparison matrix (containing a value of 1 for identities and 0 for non-identities) for bases, and the weighted comparison matrix of Gribskov and Burgess, Nucl. Acids Res., 14:6745 (1986), as described by Schwartz and Dayhoff, eds., Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, pp. 353-358 (1979); (2) a penalty of 3.0 for each gap and an additional 0.10 penalty for each symbol in each gap; and (3) no penalty for end gaps. Other programs used by those skilled in the art of sequence comparison may also be used.

[0052] Nucleic acids of the present invention also include nucleic acids which are capable of hybridizing to the base sequence of SEQ ID NO: 27 under conditions of moderately stringent conditions and functions to restore fertility, and nucleic acids which are capable of hybridizing to the base sequence of SEQ ID NO: 27 under conditions of highly stringent conditions and functions to restore fertility.

[0053] As used herein, conditions of moderate stringency can be readily determined by those having ordinary skill in the art based on, for example, the length of the DNA. The basic conditions are set forth by Sambrook et al. Molecular Cloning: A Laboratory Manual, 2nd. Vol. 1, pp. 1.101-104, Cold Spring Harbor Laboratory Press, (1989), and include use of a prewashing solution for the nitrocellulose filters 5 x SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0), hybridization conditions of about 1 x SSC to 6 x SSC at about 40°C to 60°C (or other similar hybridization solution, such as Stark's solution, in about 50% formamide at about 42°C), and washing conditions of about 60°C, 0.5 x SSC, 0.1% SDS. The hybridization temperature is about 15-20°C lower when the hybridization solution contains about 50% formamide. Conditions of high stringency can also be readily determined by the skilled artisan based on, for example, the length of the DNA. Generally, conditions of high stringency include hybridization and/or washing conditions at higher temperatures and/or lower salt concentrations than in the conditions of moderate stringency described above. For example, such conditions include hybridization conditions of 0.1 x SSC to 0.2 x SSC at about 60-65°C and/or washing conditions of 0.2 x SSC, 0.1% SDS at about 65-68°C. The skilled artisan will recognize that the temperature and wash solution

salt concentration can be adjusted as necessary according to factors such as the length of the probe.

[0054] DNAs of the present invention also include nucleic acids that differ from the base sequence of SEQ ID NO: 27 due to deletions, insertions or substitutions of one or more bases while retaining a fertility restoring function. So far as a fertility restoring function is retained, the number of bases to be deleted, inserted or substituted is not specifically limited, but preferably 1 to several thousands, more preferably 1-1000, still more preferably 1-500, even more preferably 1-200, most preferably 1-100.

[0055] Once the Rf-1 gene is further specified on the basis of the descriptions herein, it can be used by those skilled in the art after nucleic acids such as other regions than the Rf-1 gene or intron regions in the Rf-1 gene are removed. A given amino acid may be replaced, for example, by a residue having similar physiochemical characteristics. Examples of such conservative substitutions include changes from one aliphatic residue to another, such as changes from one to another of Ile, Val, Leu, or Ala; changes from one polar residue to another, such as changes between Lys and Arg, Glu and Asp, or Gln and Asn; or changes from one aromatic residue to another, such as changes from one to another of Phe, Trp, or Tyr. Other well-known conservative substitutions include e.g. changes between entire regions having similar hydrophobic characteristics. Those skilled in the art can introduce desired deletions, insertions or substitutions by well-known gene engineering techniques using e.g. site-specific mutagenesis as described in Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd edition, Cold Spring Harbor Laboratory Press, (1989).

[0056] We compared an indica variety IR24 carrying the Rf-1 gene (SEQ ID NO: 27) with japonica varieties not carrying it such as Asominori (SEQ ID NO: 28) and a Nipponbare BAC clone deposited with GenBank (Accession No. AC068923). As a result, we found that the Rf-1 region of the indica variety containing the Rf-1 gene has at least the following single bases polymorphisms (SNP).

- 1) a base corresponding to the base 1239 of SEQ ID NO: 27 is A;
- 2) a base corresponding to the base 6227 of SEQ ID NO: 27 is A;
- 3) a base corresponding to the base 20680 of SEQ ID NO: 27 is G;
- 4) a base corresponding to the base 45461 of SEQ ID NO: 27 is A;
- 5) a base corresponding to the base 49609 of SEQ ID NO: 27 is A;
- 6) a base corresponding to the base 56368 of SEQ ID NO: 27 is T;
- 7) a base corresponding to the base 57629 of SEQ ID NO: 27 is C; and
- 8) a base corresponding to the base 66267 of SEQ ID NO: 27 is G.

[0057] Thus, nucleic acids containing the Rf-1 region of the present invention preferably meet one to all of the requirements 1)-8) above.

[0058] In Example 3 below, the chromosomal organizations of recombinants proximal to the Rf-1 gene (RS1-RS2, RC1-RC8) were tested in the Rf-1 region. The results showed that a sequence determining the presence of the function of the Rf-1 gene is contained in the base sequence of bases 1239-66267 of SEQ ID NO: 27, i.e. in a region from the P4497 MboI to B56691 XbaI loci (about 65 kb) as estimated at maximum (Fig. 3). However, there is a possibility that it is important for the expression of the genetic function of the Rf-1 gene that the Rf-1 gene is partially of the indica genotype, and that the genetic function may not be significantly changed whether the remaining regions are of the japonica or indica genotype. There may be an extreme case that the coding region is completely identical and only the promoter region is different between japonica and indica, and that the promoter region and the coding region are only partially included in the region from P4497 the MboI to B56691 XbaI loci (about 65 kb). Therefore, it cannot be concluded that the common indica region above (bases 1239-66267 of SEQ ID NO: 27) completely contains the entire Rf-1 gene. However, it is thought that at least SEQ ID NO: 27 completely contains the entire Rf-1 gene for the following reasons:

- 1) the size of a gene is normally several kilobases, and rarely exceeds 10 kb;
- 2) the genomic base sequence of IR24 determined by the present invention (SEQ ID NO: 27) completely contains the common indica region above;
- 3) the 5' end of SEQ ID NO: 27 is located 1238 bp upstream of the 5' end of the common indica region above and forms a part of another gene (S12564); and
- 4) the 3' end of SEQ ID NO: 27 is located 10096 bp downstream of the 3' end of the common indica region above.

[0059] In this way, we first succeeded in restricting the region of the Rf-1 gene to 76 kb. Thus, nucleic acids containing the region of the Rf-1 gene of the present invention are extremely less likely to contain other genes proximal to the Rf-1 gene as compared with those selected with the co-dominant marker locus at a genetic distance of about 1 cM (about 300 kb) from the Rf-1 gene described in a prior documents such as Japanese Patent Public Disclosure No. 2000-139465. Moreover, they are less likely to contain other genes than those selected with the DNA marker loci S12564 Tsp509I and C1361 MwoI (at a distance of about 0.3 cM between them) described in our prior Japanese Patent Application No. 2000-247204.

[0060] We further confirmed by complementation assays that the Rf-1 gene is completely contained in especially bases 38538-54123 of the base sequence of SEQ ID NO: 27. In an embodiment of the present invention, therefore, the base sequence at least 70% identical to the base sequence of SEQ ID NO: 27 or to the base sequence of bases 38538-54123 of SEQ ID NO: 27 meets at least one of the following requirements 1) and 2):

- 1) a base corresponding to the base 45461 of SEQ ID NO: 27 is A;
- 2) a base corresponding to the base 49609 of SEQ ID NO: 27 is A.

IV. Method for restoring rice fertility

[0061] The present invention provides a method for restoring rice fertility comprising introducing a nucleic acid into rice, wherein the nucleic acid has the base sequence of SEQ ID NO.27, or has a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27, and which functions to restore fertility. The methods of the present invention may comprise introducing a nucleic acid into rice, wherein the nucleic acid has a portion of SEQ ID NO: 27, especially bases 38538-54123, preferably bases 42357-53743, more preferably bases 42132-48883 of SEQ ID NO: 27 or has a base sequence which is identical to at least 70% identical of the base sequence of bases 38538-54123, preferably bases 42357-53743, more preferably bases 42132-48883 of SEQ ID NO: 27 and, which functions to restore fertility.

[0062] In the present invention, the nucleic acid containing the locus of a fertility restorer gene (Rf-1) that can be introduced into rice can be any one of the nucleic acids described above in "III. Nucleic acids containing the Rf-1 locus". The method for introducing the nucleic acid into rice is not specifically limited but can be any known method. Nucleic acids of the present invention can be introduced by known genetic engineering techniques or crossing. Genetic engineering techniques are preferably used because inclusion of other neighboring genes can be prevented and the period for establishing a line can be shortened.

[0063] Any suitable expression system for transduction by genetic engineering techniques can be employed. Recombinant expression vectors comprise a nucleic acid containing a fertility restorer gene (Rf-1) of the invention that can be introduced into rice, operably linked to suitable transcriptional or translational regulatory base sequences, such as those derived from a mammalian, microbial, viral, or insect gene.

[0064] Examples of regulatory sequences include transcriptional promoters, operators, or enhancers, an mRNA ribosomal binding site, and appropriate sequences which control transcription and translation initiation and termination. Base sequences are operably linked to a regulatory sequence when the regulatory sequence is functionally associated with the DNA sequences. Thus, a promoter base sequence is operably linked to a DNA sequence if the promoter base sequence controls the transcription of the DNA sequence. An origin of replication that confers the ability to replicate in rice, and a selection gene by which transformants are identified, are generally incorporated into expression vectors. As for selectable markers, those commonly used can be used by standard methods. Examples are genes resistant to antibiotics such as tetracycline, ampicillin, kanamycin, neomycin, hygromycin or spectinomycin.

[0065] In addition, a sequence encoding an appropriate signal peptide (native or heterologous) can be incorporated into expression vectors. A DNA sequence for a signal peptide (secretory leader) may be fused in frame to a nucleic acid sequence of the invention so that the DNA is initially transcribed, and the mRNA translated into a fusion protein containing the signal peptide.

[0066] The present invention also provides recombinant vectors containing a gene of the present invention. Methods for integrating a DNA fragment of a gene of the present invention into a vector such as a plasmid are described in e. g. Sambrook, J. et al, Molecular Cloning, A Laboratory Manual (2nd edition), Cold Spring Harbor Laboratory, 1.53 (1989). Commercially available ligation kits (e.g. available from TAKARA) can be conveniently used. Thus obtained recombinant vectors (e.g. recombinant plasmids) are transferred into host rice cells.

[0067] Vectors can be conveniently prepared by linking a desired gene to a recombinant vector available in the art (e.g. plasmid DNA) by standard methods. Plant transforming vectors are especially useful for conferring fertility on rice using a nucleic acid fragment of the present invention. Vectors for plants are not specifically limited so far as they can express the gene of interest in plant cells to produce the protein, but preferably include pBI221, pBI121 (Clontech), and vectors derived therefrom. Especially, examples of vectors for transforming rice belonging to monocotyledons include pIG121Hm and pTOK233 (Hiei et al., Plant J., 6, 271-282 (1994)), and pSB424 (Komari et al., Plant J., 10, 165-174 (1996)).

[0068] Transgenic plants can be prepared by replacing the β -glucuronidase (GUS) gene in the above vectors with a nucleic acid fragment of the present invention to construct a plant transforming vector and transfecting it into a plant. The plant transforming vector preferably comprises at least a promoter, a start codon, a desired gene (a nucleic acid sequence of the present invention or a part thereof), a stop codon and a terminator. It may also contain a DNA encoding a signal peptide, an enhancer sequence, non-translated 5' and 3' regions of the desired gene, a selectable marker region, etc., as appropriate. Promoters and terminators are not specifically limited so far as they are functional in plant

cells, among which constitutive expression promoters include the 35S promoter initially contained in the above vectors as well as promoters for actin and ubiquitin genes.

[0069] Suitable methods for introducing a plasmid into a host cell include the use of calcium phosphate or calcium chloride/rubidium chloride, electroporation, electroinjection, chemical treatment with PEG or the like, the use of a gene gun described in Sambrook, J. et al., *Molecular Cloning, A Laboratory Manual* (2nd edition), Cold Spring Harbor Laboratory, 1.74(1989). Plant cells can be transformed by e.g. the leaf disc method [Science, 227, 129 (1985)] or electroporation [Nature, 319, 791 (1986)].

[0070] Methods for transferring a gene into a plant include the use of *Agrobacterium* (Horsch et al., Science, 227, 129 (1985); Hiei et al., Plant J., 6, 271-282(1994)), electroporation (Fromm et al., Nature, 319, 791(1986)), PEG (Paskowski et al., EMBO J., 3, 2717(1984)), microinjection (Crossway et al., Mol. Gen. Genet., 202, 179 (1986)), particle bombardment (McCabe et al., Bio/Technology, 6, 923(1988)). Methods are not specifically limited so far as they are suitable for transfecting a nucleic acid into a desired plant.

[0071] Transduction by crossing can be performed as follows, for example. First, F_1 obtained by crossing an Rf-1 donor parent and a japonica variety is backcrossed with the japonica variety. The resulting individuals are screened for those homozygous for japonica at the S12564 Tsp509I locus and heterozygous at the P4497 MboI and B53627 BstZ17I loci and further backcrossed. The resulting individuals are screened for those heterozygous at the P4497 MboI and B56691 XbaI loci and homozygous for japonica at the B53627 BstZ17I locus and further backcrossed. Subsequently, about 10 cycles of screening each backcrossed generation for individuals heterozygous at the P4497 MboI and B56691 XbaI loci and subjecting them to the subsequent backcrossing are repeated. Finally, individuals heterozygous at the P4497 MboI and B56691 XbaI loci are self-fertilized and the resulting individuals are screened for those homozygous for indica at both loci, whereby a restorer line inheriting a limited chromosomal region from the P4497 MboI to B56691 XbaI loci from the Rf-1 donor parent can be obtained.

[0072] According to the present invention, nucleic acids containing a fertility restorer gene (Rf-1) were isolated, whereby the Rf-1 gene can be introduced into a rice variety using genetic engineering techniques to establish a restorer line. The present invention succeeded in restricting the Rf-1 region to 76 kb or less. Therefore, nucleic acids containing the Rf-1 locus of the present invention are extremely less likely to contain other genes neighboring the Rf-1 gene than those of the prior art. Moreover, the entire base sequence of the region containing the Rf-1 gene was determined by the present invention. Those skilled in the art can proceed with analysis of the Rf-1 gene itself on the basis of the description herein. Thus, only the Rf-1 gene can be introduced without including any neighboring gene. This is especially important when neighboring genes bring deleterious traits. Furthermore, restorer lines can be established in a shorter period such as 1-2 years than obtained by crossing.

[0073] In complementation assays described in Examples 4-13 herein, MS Koshihikari (having BT cytoplasm and a core gene substantially identical to Koshihikari) was actually transformed by an *Agrobacterium*-mediated method using fragments from 10 clones described in Fig. 5. The results demonstrated that fertility restorer lines are established from a nucleic acid containing the base sequence of bases 38538-54123, preferably bases 42357-53743, more preferably bases 42132-48883 of SEQ ID NO: 27.

[0074] *Agrobacterium*-mediated methods for establishing rice restorer lines are described in, but not limited to, Hiei et al., Plant J., 6, pp. 271-282(1994), Komari et al., Plant J., 10, p.165-174(1996), Ditta et al., Proc. Natl. Acad. Sci. USA 77: pp. 7347-7351(1980), etc.

[0075] First, a plasmid vector containing a nucleic acid fragment of interest to be inserted is prepared. Suitable plasmid vectors include e.g. pSB11, pSB22 and the like having a plasmid map described in Komari et al., Plant J., 10, pp. 165-174 (1996), supra. Alternatively, those skilled in the art can also construct an appropriate vector by themselves on the basis of plasmid vectors such as pSB11, pSB22 described above. In the examples herein below, an intermediate vector pSB200 having a hygromycin-resistant gene cassette was prepared on the basis of pSB11, and used. Specifically, a nopaline synthase terminator (Tnos) was first fused to a ubiquitin promoter and a ubiquitin intron (Pubi-ubil). A hygromycin-resistant gene (HYG(R)) was inserted between ubil and Tnos of the resulting Pubi-ubil-Tnos complex to give a Pubi-ubil-HYG(R)-Tnos assembly. This assembly was fused to a HindIII/EcoRI fragment of pSB11 (Komari et al., supra.) to give pKY205. Linker sequences for adding restriction enzyme sites NotI, NspV, EcoRV, KpnI, SacI, EcoRI were inserted into the Hind III site upstream of Pubi of this pKY205 to give pSB200 having a hygromycin-resistant gene cassette.

[0076] Then, *E. coli* cells (e.g. DH5a, JM109, MV1184, all commercially available from e.g. TAKARA) are transformed with the recombinant vector containing the nucleic acid inserted.

[0077] Thus transformed *E. coli* cells are used for triparental mating with an *Agrobacterium* strain preferably in combination with a helper *E. coli* strain according to e.g. the method of Ditta et al. (1980). Suitable *Agrobacterium* strains include *Agrobacterium tumefaciens* strains such as LBA4404/pSB1, LBA4404/pNB1, LBA4404/pSB3, etc. They all have a plasmid map described in Komari et al., Plant J., 10, pp. 165-174 (1996), supra. and can be used by those skilled in the art by constructing a vector by themselves. Suitable helper *E. coli* strains include, but not limited to, e.g. HB101/pRK2013 (available from Clontech). A report shows that *E. coli* cells carrying pRK2073 can also be used as

helper *E. coli* though they are less common (Lemas et al., *Plasmid* 1992, 27, pp. 161-163).

[0078] Then, the *Agrobacterium* cells mated as intended are transformed into male sterility rice according to e.g. the method of Hiei et al (1994). Necessary immature rice seeds for transformation can be prepared by e.g. pollinating male sterility rice with a japonica variety.

[0079] Fertility restoration in transformed plants can be assessed by e.g. evaluating seed fertility in standing plants about one month after heading. Evaluation on standing plants means observation of plants grown in a field or the like. An alternative method is a laboratory study of grain ripening percentages in the ear.

V. Methods for discerning the presence of the Rf-1 gene

[0080] According to the present invention, it was shown that a sequence determining the presence of the function of the Rf-1 gene is located between the polymorphism-detecting marker loci P4497 Mbol and B56691 XbaI on rice chromosome 10. Moreover, complementation assays confirmed that the Rf-1 gene is completely contained in especially bases 38538-54123 of the base sequence of SEQ ID NO: 27.

[0081] Comparison of the base sequence of an indica variety carrying the Rf-1 gene (IR24) (SEQ ID NO: 27) with those of japonica varieties not carrying said gene (Asominori (SEQ ID NO: 28) and Nipponbare BAC clone AC068923) revealed the presence of polymorphisms between both varieties. As a result, it became possible to conveniently, rapidly and exactly discern whether or not a rice plant or seed under test carries the Rf-1 gene on the basis of polymorphisms in base sequence in regions neighboring the Rf-1 gene.

[0082] Therefore, the present invention also provides a method for discerning whether or not a subject rice individual or a seed thereof has the Rf-1 gene or not, wherein the method utilizing a fact that a sequence determining the presence of the function of the Rf-1 gene positions between the polymorphism detection marker loci P4497 Mbol and B56691 XbaI on rice chromosome 10.

[0083] Polymorphisms can be detected by any known method. For example, known methods include assays for restriction fragment length polymorphisms (RFLPs); direct determination by sequencing; cutting a genomic DNA with a 8-base recognizing restriction enzyme, and then radioactively labeling the ends and further cutting the labeled digest with 6-base and 4-bases recognizing restriction enzyme and then developing the digest by two-dimensional electrophoresis (RLGS, Restriction Landmark Genome Scanning); etc. AFLP analysis (amplified fragment length polymorphism; P. Vos et al., *Nucleic Acids Res.* Vol. 23, pp. 4407-4414 (1995)) has also been developed wherein RFLP is amplified/detected by polymerase chain reaction (PCR).

[0084] For example, conventional methods involved detecting RFLPs via PCR amplification (conversion of RFLP markers into PCR markers) or detecting polymorphisms in microsatellites via PCR amplification (microsatellite markers) as illustrated below.

Conversion of RFLP markers into PCR markers

[0085]

A. PCR markers based on polymorphisms in genomic regions corresponding to RFLP probes (D.E. Harry, B. Temesgen, D.B. Neale; *Codominant PCR-based markers for Pinus taeda developed from mapped cDNA clones*, *Theor. Appl. Genet.* (1998) 97: pp. 327-336). After performing genomic PCR using primers designed for an RFLP marker probe sequence ("RFLP" is a polymorphism observed by Southern analysis using a DNA fragment as a probe. The base sequence of the DNA fragment used as a probe is called "RFLP marker probe sequence"), a PCR marker can be prepared by either of the following two procedures. A first procedure involves treating the products with a series of restriction enzymes to search for a restriction enzyme causing a fragment length polymorphism, and a second procedure involves searching for a polymorphism by varietal comparison of the base sequences of the products and preparing a PCR marker based on the polymorphism.

B. PCR markers based on identification of RFLP-causing sites. A PCR marker can be obtained by identifying an RFLP-causing site (a restriction enzyme recognition site carried by only one of two varieties compared) present in or near (normally within several kbs) an RFLP marker probe sequence.

Microsatellite markers

[0086] Microsatellites are repeat sequences of about 2 to 4 bases such as (CA)_n that are present in great numbers in genomes. If a varietal polymorphism occurs in repetition number, a polymorphism can be observed in PCR product length by PCR using primers designed in adjacent regions, whereby the DNA polymorphism can be detected. Markers for detecting polymorphisms using microsatellites are called microsatellite markers (O. Parnaud, X. Chen, S.R. McCouch, *Mol. Gen. Genet.* (1996) 252: pp. 597-607).

[0087] Methods for detecting polymorphisms in the present invention are not specifically limited. From the viewpoint of efficiency and convenience, PCR-RFLP is preferred, which is a combination of PCR and RFLP to identify polymorphisms from their restriction enzyme cleavage patterns in cases where they exist among variety lines at restriction enzyme recognition sites in the sequences of DNA fragments amplified by PCR. PCR-RFLP is also called CAPS (cleaved amplified polymorphic sequence). If any suitable restriction enzyme recognition site is not present in a region showing polymorphisms, a modified CAPS called dCAPS (derived cleaved amplified polymorphic sequence) can also be used wherein restriction enzyme sites are introduced during PCR (Michaels, S.D. and Amasino, R.M. (1998), The Plant Journal 14(3) 381-385; A. Konieczny et al., (1993), Plant J. 4(2) pp. 403-410; Neff, M.M., Neff, J.D., Chory, J. and Pepper, A.E. (1998), The Plant Journal 14(3) 387-392). These methods are explained in more detail below.

CAPS, dCAPS

[0088] The method for discerning of the present invention comprise, but not limited to:

- i) preparing a pair of primers based on the base sequences of a site showing a polymorphism in the base sequences between indica and japonica varieties at the Rf-1 locus and its adjacent regions to amplify said base sequences;
- ii) performing nucleic acid amplification reaction(s) using the genomic DNA of the subject rice individual or the seed thereof as a template; and
- iii) discerning whether or not the subject rice individual or the seed thereof has the Rf-1 gene based on the polymorphism found in the nucleic acid amplification product.

[0089] The step of preparing a primer pair in i) preferably comprises any of the following means:

- a) when a change containing a deleted region exists in the polymorphism in the nucleic acid amplification product, preparing a pair of primers for nucleic acid amplification to flank the deleted region to form a marker for detecting the polymorphism;
- b) when a base change causing a difference in restriction enzyme recognition exists in the polymorphism in the nucleic acid amplification product, preparing a pair of primers for nucleic acid amplification to flank the base change site to form a marker for detecting the polymorphism; or
- c) when a base change causing no difference in restriction enzyme recognition exists in the polymorphism in the nucleic acid amplification product, preparing a pair of primers for introducing a mismatch, wherein pair of primers contain the base change site and alters a region containing the base change site into a base sequence causing a difference in restriction enzyme recognition in the nucleic acid amplification product to form a marker for detecting the polymorphism.

[0090] Suitable polymorphic sites for discerning the presence of the Rf-1 gene in the present invention can be appropriately selected so that a polymorphism detecting marker can be prepared as described below on the basis of comparison of, but not limited to, the base sequence of an indica variety carrying the Rf-1 gene (IR24) (SEQ ID NO: 27) with those of japonica varieties not carrying said gene (Asominori (SEQ ID NO: 28) and Nipponbare BAC clone AC068923).

[0091] If the polymorphism found causes a difference in restriction enzyme recognition, for example, a pair of primers for nucleic acid amplification are prepared to flank the polymorphic site and used for detecting the polymorphism. Primers are preferably designed not to be specific for highly repeated sequences to avoid undesired products. If the polymorphism found does not cause a difference in restriction enzyme recognition, markers can be prepared by applying the dCAPS method described above. Primers for dCAPS markers are preferably designed not to be specific for repeat sequences and to provide a product length of preferably 50-300 bases, more preferably about 100 bases to ease identification of polymorphisms.

[0092] If the polymorphism found involves a microsatellite, nucleic acid amplification primers are prepared to flank the microsatellite and used to detect the polymorphism. Again, the primers are preferably designed not to be specific for repeat sequences.

1) Nucleic acid amplification

[0093] In the present invention, a pair of primers are preferably prepared for amplifying adjacent regions containing polymorphisms on the basis of the determined base sequence of the nucleic acid of a subject rice individual or seed at the Rf-1 locus. The primer pair is used to perform a nucleic acid amplification reaction with the genomic DNA of the subject rice individual or seed as a template. The nucleic acid amplification reaction is preferably polymerase chain reaction (PCR) (Saiki et al., 1985, Science 230, pp. 1350-1354).

[0094] The pair of primers for nucleic acid amplification can be prepared by any known method on the basis of the base sequence of a polymorphic site and adjacent regions thereto. Specifically, a primer pair can be prepared on the basis of the base sequence of a polymorphic site and adjacent regions thereto by a process comprising generating a single-stranded DNA having the same base sequence as the base sequence of the polymorphic site and adjacent regions thereto or a base sequence complementary to said regions or, if necessary, generating the single-stranded DNA containing a modification without affecting the binding specificity to the base sequence of the polymorphic site and adjacent regions thereto provided that the following conditions are satisfied:

- 1) the length of each primer should be 15-30 bases;
- 2) the proportion of G+C in the base sequence of each primer should be 30-70%;
- 3) the distribution of A, T, G and C in the base sequence of each primer should not be partially largely uneven;
- 4) the length of the nucleic acid amplification product amplified by the primer pair should be 50-3000 bases, preferably 50-300 bases; and
- 5) any complementary sequence segment should not occur with the base sequence of each primer or between the base sequences of the primers.

[0095] As used herein, the "adjacent regions" to a polymorphic site mean that an area containing both of a polymorphic site and adjacent regions thereto is within a distance suitable for nucleic acid amplification, preferably PCR. The adjacent regions amplified preferably have a length within the range of, but not limited to, about 50 bases to about 3000 bases, more preferably about 50 bases to about 2000 bases. To facilitate identification of polymorphisms, the product length is preferably 50-300 bases, more preferably about 100 bases. The adjacent regions preferably have a length within the range of, but not limited to, about 0 to about 3000 bases, more preferably about 0 to about 2000 bases, still more preferably about 0 to about 1000 bases on the 5' or 3' side of a polymorphic site.

[0096] Procedures and conditions for the nucleic acid amplification reaction are not specifically limited and are well known to those skilled in the art. Appropriate conditions can be applied by those skilled in the art depending on various factors such as the base sequence of the polymorphic site and adjacent regions thereto, the base sequence and length of the primer pair, etc. Generally, the nucleic acid amplification reaction can be performed under more stringent conditions (annealing reaction and nucleic acid elongation reaction at higher temperatures and less cycles) as the primer pair is longer or the proportion of G+C is higher or the distribution of A, T, G and C is even. The use of more stringent conditions allows an amplification reaction with higher specificity.

[0097] The amplification reaction can be performed under conditions of, but not limited to, one cycle of 94°C for 2 min, 30 cycles of 94°C for 1 min, 58 °C for 1 min and 72°C for 2 min, and finally one cycle of 72°C for 2 min using 50 ng of a genomic DNA as a template, 200 µM of each dNTP and 5 U of ExTaq™ (TAKARA). The reaction can also be performed under conditions of one cycle of 94°C for 2 min, 30 cycles of 94°C for 1 min, 58°C for 1 min and 72°C for 1 min, and finally one cycle of 72°C for 2 min. In another embodiment, the reaction can also be performed under conditions of one cycle of 94°C for 2 min, 35 cycles of 94°C for 30 sec, 58°C for 30 sec and 72°C for 30 sec, and finally one cycle of 72°C for 2 min.

[0098] The subject rice (test rice) genomic DNA used as a template for PCR can be easily extracted from individuals or seeds by the method of Edwards et al. (Nucleic Acids Res. 8(6):1349, 1991). More preferably, DNA purified by standard techniques is used. An especially preferred extraction method is the CTAB method (Murray, M.G. et al., Nucleic Acids Res. 8(19):4321-5, 1980). The DNA is preferably used as a template for PCR at a final concentration of 0.5 ng/µL.

2) Preparation of markers for detecting polymorphisms

[0099] After examining whether or not a polymorphism is detected in the amplification product by the nucleic acid amplification reaction with a pair of primers, a marker for detecting the polymorphism is prepared on the basis of the polymorphism found. Non-limiting examples of polymorphisms that can be detected in the amplification product are as follows.

a) A change containing a deleted region exists in the polymorphism in the nucleic acid amplification product.

[0100] In this case, a pair of primers for nucleic acid amplification are prepared to flank the deleted region to form a marker for detecting the polymorphism. If the deleted region has a sufficient size, the polymorphism can be detected from the difference in mobility by electrophoresing the amplification product on an agarose gel or an acrylamide gel, for example. The polymorphism can be detected when the difference in base pair numbers is about 5% or more in the case of agarose gel electrophoresis or when the difference in length is about 1 base or more in the case of sequencing acrylamide gel electrophoresis, for example. Alternatively, the polymorphism can be detected by hybridizing the nucleic

acid amplification product using an oligobase or a DNA fragment having a complementary sequence to the base sequence excluding the deleted region as an analytical probe. Alternatively, the polymorphism can be confirmed by determining the base sequence of the amplification product, if desired. Known techniques for electrophoresis of nucleic acids, hybridization, sequencing and the like can be used as appropriate by those skilled in the art. In this case, the difference in the length of the amplification product directly reflects the polymorphism and markers for detecting polymorphisms on this basis are called ALP (amplicon length polymorphism) markers.

b) A base change causing a difference in restriction enzyme recognition exists in the polymorphism in the nucleic acid amplification product.

[0101] In this case, a pair of primers for nucleic acid amplification are prepared to flank the base change site to form a marker for detecting the polymorphism. In this case, a base change causing a difference in restriction enzyme recognition occurs in the polymorphism of the nucleic acid amplification product, i.e. the nucleic acid amplification product may be cleaved or not with one or more specific restriction enzymes. Thus, the amplification product can be treated with the restriction enzymes and electrophoresed on e.g. an agarose gel to detect the polymorphism from the difference in mobility. The polymorphism can be confirmed by determining the base sequence of the amplification product, if desired.

[0102] In this case, the difference in the length of the restriction fragment of the amplification product by PCR or the like reflects the polymorphism and markers for detecting polymorphisms on this basis are called CAPS markers or PCR-RFLP markers (A. Konieczny et al., supra.)

[0103] This is exemplified by primer pairs P4497 MboI, P23945 MboI, P41030 TaqI, P45177 BstUI, B59066 BsaJI and B56691 XbaI in Example 1 below. Even if the polymorphism can be detected by the length of the nucleic acid amplification product as described in a) above, the polymorphism can be more easily detected by combination with restriction enzyme treatment.

c) A base change causing no difference in restriction enzyme recognition exists in the polymorphism in the nucleic acid amplification product.

[0104] In this case, a pair of primers for introducing a mismatch are prepared that contains the base change site and alters a region containing the base change site into a base sequence causing a difference in restriction enzyme recognition in the nucleic acid amplification product to form a marker for detecting the polymorphism.

[0105] Specifically, a pair of primers based on the base sequences of regions naturally proximal to the Rf-1 gene cause a polymorphism in the nucleic acid amplification product but no difference in restriction enzyme recognition, and therefore, a mismatch is introduced into one or both of the primers to alter a region containing the base change site (polymorphism) into a base sequence causing a difference in restriction enzyme recognition in the nucleic acid amplification product. For example, the method described in Mikaelian et al., Nucl. Acids. Res. 20:376.1992 can be used as a standard technique for substituting, deleting or adding a specific base by PCR-mediated site-specific mutagenesis. The amplification product using the mismatch-introducing primers as a marker for detecting the polymorphism may be cleaved or not with one or more specific restriction enzymes because it has a difference in restriction enzyme recognition at the mismatch-introducing site. Therefore, the amplification product can be treated with the restriction enzymes and electrophoresed on e.g. an agarose gel to detect the polymorphism from the difference in mobility, as described in b) above.

[0106] The introduction of a mismatch must not affect not only the binding of the primers to a target plant genome but also the polymorphic base change. The polymorphic base change is used to introduce a mismatch near it so that a difference in restriction enzyme recognition occurs by a combination of both base change and mismatch. Methods for introducing such a mismatch are known to those skilled in the art and described in detail in Michaels, S.D. and Amasino, R.M. (1998), Neff, M.M., Neff, J.D., Chory, J. and Pepper, A.E. (1998), for example.

[0107] Markers in this case are improved CAPS markers described in b) above and called dCAPS (derived CAPS) markers. This is exemplified by P9493 BslI in Example 3 below.

[0108] If there are many extra restriction sites unrelated to varietal polymorphisms in the case of b) or c) above, it may be difficult to discern any difference in restriction site recognition based on polymorphisms. In this case, a mismatch may be introduced into a primer as appropriate to abolish unnecessary restriction sites. For example, a mismatch was introduced into the R-primer to abolish the MspI site unrelated to polymorphisms in B60304 MspI in Example 3.

[0109] Although the invention is not limited to any specific method, CAPS or dCAPS methods have several advantages over other RFLP methods. Specifically, analyses can be made with smaller amounts of samples than in RFLP, for example. Another advantage is that the time and labor required for analyses can be greatly reduced. Polymorphisms detected with PCR markers can be visualized by agarose gel electrophoresis that is easier than acrylamide gel electrophoresis used for microsatellite markers.

Preferred embodiments of the discerning method of the present invention

[0110] Preferred embodiments of the method for discerning whether or not a subject rice has the Rf-1 gene are described below for illustrative purposes. In the examples herein, it was found that the base sequence of an indica variety IR24 carrying the Rf-1 gene (SEQ ID NO: 27) has at least the following polymorphisms 1)-8) as compared with corresponding regions of japonica varieties:

- 1) a base corresponding to the base 1239 of SEQ ID NO: 27 is A;
- 2) a base corresponding to the base 6227 of SEQ ID NO: 27 is A;
- 3) a base corresponding to the base 20680 of SEQ ID NO: 27 is G;
- 4) a base corresponding to the base 45461 of SEQ ID NO: 27 is A;
- 5) a base corresponding to the base 49609 of SEQ ID NO: 27 is A;
- 6) a base corresponding to the base 56368 of SEQ ID NO: 27 is T;
- 7) a base corresponding to the base 57629 of SEQ ID NO: 27 is C; and
- 8) a base corresponding to the base 66267 of SEQ ID NO: 27 is G.

[0111] In preferred embodiments of the present invention, therefore, the subject rice individual or seed is judged as carrying the Rf-1 gene when one to all of the requirements 1)-8) above are met.

[0112] We further verified that a region essential for the expression of the function of the Rf-1 gene is contained in especially bases 38538-54123, preferably bases 42357-53743, more preferably bases 42132-48883 in the base sequence of SEQ ID NO: 27. In an embodiment of the present invention, therefore, the subject rice individual or seed is determined to have the Rf-1 gene in the case that the nucleic acid having a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO: 27 or of the base sequence of bases 38538-54123 of SEQ ID NO: 27, meets at least one of the following requirements 1) and 2):

- 1) a base corresponding to the base 45461 of SEQ ID NO: 27 is A; and
- 2) a base corresponding to the base 49609 of SEQ ID NO: 27 is A.

[0113] Known polymorphism detecting methods can be used to determine whether or not the above requirements are met. The base sequence of adjacent regions containing said sequence can also be directly determined. However, CAPS or dCAPS methods described above are preferably used because they are rapid and convenient. CAPS or dCAPS methods can be performed by a protocol comprising, for example:

i) preparing a pair of primers based on a base sequence of adjacent regions including any one of the following base;

- 1) a base corresponding to the base 1239 of SEQ ID NO: 27;
- 2) a base corresponding to the base 6227 of SEQ ID NO: 27;
- 3) a base corresponding to the base 20680 of SEQ ID NO: 27;
- 4) a base corresponding to the base 45461 of SEQ ID NO: 27;
- 5) a base corresponding to the base 49609 of SEQ ID NO: 27;
- 6) a base corresponding to the base 56368 of SEQ ID NO: 27;
- 7) a base corresponding to the base 57629 of SEQ ID NO: 27; and
- 8) a base corresponding to the base 66267 of SEQ ID NO: 27 is G.

to amplify both the base of the above and adjacent regions thereto;

ii) performing nucleic acid amplification reaction(s) using the genome DNA of the subject rice individual or the seed thereof as a template; and

iii) discerning the presence of the Rf-1 in the subject rice individual or the seed thereof based on polymorphism found in said nucleic acid amplification product.

[0114] The detection of polymorphisms in the nucleic acid amplification product is performed by, but not limited to, discerning the subject rice individual or seed to have the Rf-1 gene when one to all of the requirements 1)-8) below are met:

- 1) a region including a base corresponding to the base 1239 of SEQ ID NO: 27 does not have any MboI recognition sequence;
- 2) a region including a base corresponding to the base 6227 of SEQ ID NO: 27 does not have any BstI recognition sequence;

- 3) a region including a base corresponding to the base 20680 of SEQ ID NO: 27 does not have any MboI recognition sequence;
- 4) a region including a base corresponding to the base 45461 of SEQ ID NO: 27 does not have any TaqI recognition sequence;
- 5) a region including a base corresponding to the base 49609 of SEQ ID NO: 27 does not have any BstUI recognition sequence;
- 6) a region including a base corresponding to the base 56368 of SEQ ID NO: 27 does not have any MspI recognition sequence;
- 7) a region including a base corresponding to the base 57629 of SEQ ID NO: 27 does not have any BsaJI recognition sequence; and
- 8) a region including a base corresponding to the base 66267 of SEQ ID NO: 27 does not have any XbaI recognition sequence.

[0115] However, the present invention is not limited to the restriction enzymes above so far as each polymorphism in the specific regions 1)-8) above can be detected.

[0116] Preferably, identification methods of the present invention comprise:

i) preparing a pair of primers based on a base sequence of adjacent regions including any one of the following base;

- 1) a base corresponding to the base 45461; or
- 2) a base corresponding to the base 49609; to amplify both the base of the above and adjacent regions thereto;

ii) performing nucleic acid amplification reaction(s) using the genome DNA of the subject rice individual or the seed thereof as a template; and

iii) discerning the presence of the Rf-1 in the subject rice individual or the seed thereof based on polymorphism found in said nucleic acid amplification product. The subject rice individual or seed thereof is determined to have the Rf-1 gene in step iii), although not limited to, when at least one of the following requirements 1) and 2) is met:

- 1) a region including a base corresponding to the base 45461 of SEQ ID NO: 27 does not have any TaqI recognition sequence;
- 2) a region including a base corresponding to the base 49609 of SEQ ID NO: 27 does not have any BstUI recognition sequence.

[0117] Primer pairs used for the amplification reaction can be appropriately selected by those skilled in the art to preferably satisfy the conditions above on the basis of the base sequence of SEQ ID NO: 27. Preferably, any primer pair having a base sequence selected from the group consisting of SEQ ID NOS: 39 and 40, SEQ ID NOS: 41 and 42, SEQ ID NOS: 43 and 44, SEQ ID NOS: 45 and 46, SEQ ID NOS: 47 and 48, SEQ ID NOS: 49 and 50, SEQ ID NOS: 51 and 52, and SEQ ID NOS: 53 and 54 is used. More preferably, the primer pair is selected from the group consisting of SEQ ID NOS: 45 and 46, and SEQ ID NOS: 47 and 48. If necessary, the sequences of the above primer pairs containing substitutions, deletions or additions while retaining the binding specificity for the base sequence of the polymorphic site and adjacent regions thereto can also be used as primers.

[0118] To examine the resulting PCR product for restriction fragment length polymorphisms, it is cleaved with restriction enzymes corresponding to the restriction sites present in PCR markers. Such cleavage is accomplished by incubation for several hours to a day at the recommended reaction temperature for the restriction enzymes used. The PCR amplified sample cleaved with the restriction enzymes can be analyzed by electrophoresis on an about 0.7% - 2% agarose gel or an about 3% MetaPhor™ agarose gel. The gel is visualized under UV light in ethidium bromide, for example.

[0119] In the most preferred embodiments of the present invention, restriction enzyme cleavage patterns show the bands as shown in Table 2 below on the visualized gel depending on the primer pair used.

Table 2

	Approximate size (bp) of detected band
Amplified with P4497 MboI (SEQ ID NOS: 39 and 40) Restriction enzyme MboI	730
Test rice genome having the Rf-1 gene (homozygous) no	385, 345

Table 2 (continued)

		Approximate size (bp) of detected band
5	Amplified with P4497 MboI (SEQ ID NOS: 39 and 40) Restriction enzyme MboI	730
	Amplified with P9493 BslI (SEQ ID NOS: 41 and 42) Restriction enzyme BslI	126
10	Test rice genome having the Rf-1 gene (homozygous)	
	no	100, 26
	Amplified with P23945 MboI (SEQ ID NOS: 43 and 44) Restriction enzyme MboI	
15	Test rice genome having the Rf-1 gene (homozygous)	160, 100
	no	260
	Amplified with P41030 TaqI (SEQ ID NOS: 45 and 46) Restriction enzyme TaqI	
20	Test rice genome having the Rf-1 gene (homozygous)	280
	no	90, 190
	Amplified with P45177 BstUI (SEQ ID NOS: 47 and Restriction enzyme BstUI	48)
25	Test rice genome having the Rf-1 gene (homozygous)	20,65,730
	no	20,65,175,555
	Amplified with B60304 MspI (SEQ ID NOS: 49 and 50) Restriction enzyme MspI	
30	Test rice genome having the Rf-1 gene (homozygous)	330
	no	220, 110
	Amplified with B59066 BsaJI (SEQ ID NOS: 51 and Restriction enzyme BsaJI 52)	
35	Test rice genome having the Rf-1 gene (homozygous)	420
	no	65, 355
	Amplified with B56691 XbaI (SEQ ID NOS: 53 and 54) Restriction enzyme XbaI	
40	Test rice genome having the Rf-1 gene (homozygous)	670
	no	140, 530

[0120] In Example 3 below, recombinants proximal to the Rf-1 gene having pollen fertility (RS1-RS2, RC1-RC8) were tested for the chromosomal organization of the Rf-1 region using 14 polymorphic markers including the 8 primer pairs described above. As a result, it was confirmed that all the plants carry the Rf-1 gene derived from the indica variety between P9493 BslI and 59066 BsaJI. This result showed that recombinant pollens having the chromosomal organization as shown in Fig. 3 have pollen fertility, i.e. the Rf-1 gene is functional in these pollens. This means that a sequence determining the presence of the function of the Rf-1 gene is included in the indica region common to these recombinant pollens, i.e. in a region from the P4497 MboI to B56691 XbaI loci (about 65 kb) as estimated at maximum.

[0121] In the present invention, chromosomal walking was started on the presumption that the S12564 Tsp509I locus should be very proximal to the Rf-1 locus as judged from the frequency of appearance of individuals by crossing. In fact, the genetic distance between both loci has been calculated to be about 0.04 cM as the result of the high-precision segregation analysis of the present invention. Even one of markers known to be most closely linked to the Rf-1 locus as described in Japanese Patent Public Disclosure No. 2000-139465 is reported to have a genetic distance of 1 cM from the Rf-1 locus. Considering that 1 cM is estimated to be equivalent to 300 kb on average in rice, a considerable time should be required to restrict the Rf-1 gene region if chromosomal walking were started from the marker described in Japanese Patent Public Disclosure No. 2000-139465.

VI. Method for inhibiting the function of Rf-1 gene to restore fertility

[0122] According to the present invention, the nucleic acid containing the locus of a fertility restorer gene (Rf-1) including the nucleic acids which function to restore fertility was isolated. The entire base sequence thereof was determined, whereby the fertility restoring function of the Rf-1 gene can be controlled by genetic engineering techniques. Thus, the present invention further provides a method for inhibiting the function of Rf-1 to restore fertility.

[0123] A method for inhibiting the function of the Rf-1 gene to restore fertility according to one embodiment of the present invention comprises introducing an antisense having at least 100 continuous bases in length, and having a base sequence complementary to a nucleic acid having the base sequence of SEQ ID NO.27, or to a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27, and which functions to restore fertility.

[0124] In an embodiment, the method for inhibiting the function of the Rf-1 gene to restore fertility according to the present invention comprises introducing an antisense having at least 100 continuous bases in length, and being selected from base sequences complementary to a nucleic acid having the base sequence of bases 38538-54123, preferably bases 42357-53743, more preferably bases 42132-48883 of SEQ ID NO: 27, or to a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of bases 38538-54123, preferably bases 42357-53743, more preferably bases 42132-48883 of SEQ ID NO: 27 and, which functions to restore fertility.

[0125] The antisense has a length of at least 100 bases or more, more preferably 500 bases or more, most preferably 1000 bases or more. From the viewpoint of technical convenience of introduction, it preferably has a length of 10000 bases or less, more preferably 5000 bases or less. The antisense can be synthesized by known methods. The antisense can be introduced into rice by known methods as described in e.g. Terada et al. (Plant Cell Physiol. 2000 Jul, 41(7), pp. 881-888).

[0126] It is also anticipated that Rf-1 disrupted lines can be established by screening variant lines containing a transposable element such as, but not limited to, Tos17 (Hirochika H. et al. 1996, Proc. Natl. Acad. Sci. USA 93, pp. 7783-7788) for a line containing the transposable element in the base sequence of SEQ ID NO: 27. In plants, gene disruption by homologous recombination has been studied. It may also be possible to inhibit fertility restoring function by establishing such a line in which the Rf-1 gene has been replaced by a variant Rf-1 gene using a nucleic acid having the base sequence of SEQ ID NO.27, or a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27.

References[0127]

1. Fukuta et al. 1992, Jpn J. Breed. 42 (supl.1) p.164-165.
2. Japanese Patent Public Disclosure No. HEI7(1995)-222588.
3. Japanese Patent Public Disclosure No. HEI9(1997)-313187.
4. Japanese Patent Public Disclosure No. 2000-139465.
5. Harushima et al. 1998, Genetics 148 p.479-494.
6. Michaels and Amasino 1998, The Plant Journal 14(3) p.381-385.
7. Neff et al. 1998, The plant Journal 14(3) p.387-392.
8. D.E. Harry, et al., Theor Appl Genet (1998) 97:p.327-336.
9. Hiei et al., Plant Journal (1994),6(2),p.272-282.
10. Komari et al., Plant Journal (1996) 10, p.165-174.
11. Ditta et al., Proc.Natl.Acad.Sci. USA (1980), 77: p.7347-7351,
12. P. Vos et al., Nucleic Acids Res. Vol.23, p. 4407-4414 (1995).
13. O.Parnaud,X. et al, Mol.Gen.Genet.(1996) 252:p.597-607.
14. A.Konieczny et al.,(1993),Plant J.4(2)p.403-410.
15. Edwards et al.,Nucleic Acids Res. 8(6): 1349, 1991.
16. Murray M.G. et al., Nucleic Acids Res. 8(19):4321-5, 1980.
17. Terada et al., Plant Cell Physiol. 2000 Jul, 41(7), p.881-888.
18. Hirochika H. et al. 1996, Proc.Natl.Acad.Sci.USA 93, p.7783-7788.

Examples

[0128] The following examples further illustrate the present invention but are not intended to limit the technical scope of the invention. Those skilled in the art can readily add modifications/changes to the present invention on the basis of the description of the specification, and those modifications/changes are included in the technical scope of the

present invention.

Reference examples

5 [0129] The following reference examples are based on the examples described in our prior application (Japanese Patent Application No. 2000-247204 filed August 17, 2000).

Reference example 1: Conversion of RFLP markers around Rf-1 gene to PCR markers

10 [0130] In this reference example, nine RFLP markers (i.e., R1877, G291, R2303, S12564, C1361, S10019, G4003, S10602 and G2155) around the locus of Rf-1 gene were converted to PCR markers.

(1) Materials and methods

15 [0131] The following nine RFLP markers, R1877, G291, R2303, S12564, C1361, S10019, G4003, S10602 and G2155, were purchased from the National Institute of Agrobiological Sciences, the Ministry of Agriculture, Forestry and Fisheries of Japan. After determining the base sequences of the inserts in the vectors, experiments were conducted according to the following procedures. Among rice varieties herein, Asominori belongs to japonica, and IR24 belongs to indica.

(2) Preparation of Asominori genomic library

20 [0132] Total DNA was extracted from green leaves of Asominori by the CTAB method. After partial digestion with MboI, the DNA was fractionated according to size by NaCl density gradient centrifugation (6-20% linear gradient, 20°C, 37,000 rpm, 4 hr, total volume = 12 mL). A portion of each fraction (about 0.5 mL) was subjected to electrophoresis and fractions containing 15-20 kb DNA were collected and purified. A library was constructed using Lambda DASH II (Stratagene) as a vector in accordance with the attached protocol. Giga Pack III Gold (Stratagene) was used for packaging. After packaging, 500 µL of SM Buffer and 20 µL of chloroform were added. After centrifugation, 20 µL of chloroform was added to the supernatant to make a library solution.

30 [0133] XL-1 Blue MRA (P2) was infected with 5 µL of a 50-fold dilution of the library solution, whereupon 83 plaques were formed. This corresponded to 4.15×10^5 pfu per library, and hence, it was calculated that the plaques covered 8.3×10^9 bp assuming that the average length of the inserted fragments was 20 kb. The library was therefore considered to have an adequate size for the rice genome (4×10^8 bp).

35 (3) Isolation of genomic clones containing R1877-, C1361- and G4003-marker regions.

40 [0134] As for C1361 and G4003, plasmids containing the RFLP marker probe were isolated and subjected to restriction enzyme treatment and electrophoresis to separate the RFLP marker probe portion; the desired DNA was recovered on a DNA recovery filter (Takara SUPREC-01). As for R1877, primers were designed that were specific to both ends of the marker probe and PCR was performed with the total DNA of Asominori used as a template; the amplification products were electrophoresed and recovered by the method described above. The recovered DNA was labelled with a Rediprime DNA Labelling System (Amersham Pharmacia) to prepare a probe for screening the library. PCR was performed in the usual manner (this also applies to the following description).

45 [0135] Screening of the library was performed in the usual manner after blotting the plaques onto Hybond-N+ (Amersham Pharmacia). After primary screening, areas of positive plaques were individually punched out, suspended in SM buffer and subjected to the second round of screening. After the second screening, the positive plaques were punched out and subjected to the third round of screening to isolate a single plaque.

50 [0136] The isolated plaque of interest was suspended in SM buffer and primary multiplication of the phage was performed by the plate lysate method. The resulting phage-enriched solution was subjected to secondary multiplication by shake culture and the phage DNA was purified with Lambda starter kit (QIAGEN).

[0137] For each marker, primary screening was conducted on eight plates. A 10 µL aliquot of the library solution was employed per plate. After the primary, second and third rounds of screening, four genomic clones in association with R1877 were isolated and three were isolated in association with each of C1361 and G4003.

55 (4) Conversion of R1877 to PCR marker

[0138] The isolated genomic clones were analyzed to identify the causative site of RFLP, or the EcoRI site that exists in IR24 (indica rice) but not in Asominori (japonica rice), thereby converting R1877 to a PCR marker.

[0139] Specifically, the four isolated clones were subjected to the following analyses. First, T3 and T7 primers were used to determine the base sequences at both ends of the insert in each clone. Then, primers extending outwardly from both ends of the marker probe were designed. They were combined with T3 and T7 primers to give a combination of four primers in total, and employed in PCR with each clone used as the template.

[0140] In a separate step, each clone was digested with NotI and EcoRI, and electrophoresed to estimate the insert size and the length of each EcoRI fragment.

[0141] These analyses revealed the relative positions of the individual clones. In RFLP analysis, marker probe R1877 was reported to detect an EcoRI fragment of 20 kb in Nipponbare (japonica rice) and one of 6.4 kb in Kasalath (indica rice) (<ftp://ftp.staff.or.jp/pub/geneticmap98/parentsouthern/chrl0/R1877.JPG>). This fact, taken together with the results of analysis described above, gave a putative position for the EcoRI site that existed in IR24 but not in Asominori. Hence, a primer combination (SEQ ID NO:1 x SEQ ID NO:2) that was designed to amplify the nearby region was employed to perform genomic PCR over 30 cycles, each cycle consisting of 94°C x 1 min, 58°C x 1 min and 72°C x 2 min. The PCR product was treated with EcoRI and subjected to electrophoresis on 0.7% agarose gel.

[0142] As a result, the expected polymorphisms were observed between Asominori and IR24. By treatment with EcoRI, the PCR product (~3200 bp) was cleaved to yield 1500 bp and 1700 bp fragments in IR24 but not in Asominori. Mapping of the marker was made with an RIL (recombinant inbred line) of Asominori-IR24 with the results that the PCR marker was located in the same region as that of RFLP marker locus R1877, thereby confirming the conversion of RFLP marker R1877 to a PCR marker, which was named R1877 EcoRI in the present invention.

(5) Conversion of G4003 to PCR marker

[0143] The isolated genomic clones were analyzed to identify the causative site of RFLP, or the HindIII site that existed in Asominori but not in IR24, thereby converting G4003 to a PCR marker.

[0144] By performing analyses similar to those employed for R1877, the relative positions of the three isolated clones were revealed. In RFLP analysis, marker probe G4003 was reported to detect a HindIII fragment of 3 kb in Nipponbare (japonica rice) and one of 10 kb in Kasalath (indica rice) (<ftp://ftp.staff.or.jp/pub/geneticmap98/parentsouthern/chrl0/R1877.JPG>). This report, taken together with the analyses described above, led to a temporary conclusion that the HindIII site that existed in Asominori but not in IR24 would be at either one of two candidate sites. Hence, a primer combination (SEQ ID NOS: 3 and 4) that was designed to amplify the area in the neighborhood of each HindIII site was employed to perform genomic PCR over 35 cycles, each cycle consisting of 94°C x 30 sec, 58°C x 30 sec and 72°C x 30 sec. The PCR product was treated with HindIII and subjected to electrophoresis on 2% agarose gel. As a result, the HindIII site within the marker probe was found to have polymorphisms. By treatment with HindIII, the PCR product (362 bp) was cleaved to yield a 95 bp fragment and a 267 bp fragment in Asominori but not in IR24. Mapping of the site demonstrated the conversion of RFLP marker G4003 to a PCR marker, which was named G4003 HindIII (SEQ ID NO:19) in the present invention.

(6) Conversion of C1361 to PCR marker

[0145] Primers were designed on the basis of the base sequence information of the isolated genomic clones. PCR was performed with the total DNAs of Asominori and IR24 being used as a template and the PCR product was recovered by known methods after electrophoresis. Using the recovered DNA as a template, the inventors analyzed the base sequence of each of the rice varieties with ABI Model 310 in search of mutations that would cause polymorphisms.

[0146] By performing analyses similar to those employed for R1877, approximate relative positions of the three isolated clones could be established. As it turned out, however, regions around the C1361 marker would be difficult to amplify by PCR or determine their base sequences, and hence, it would not be easy to identify the causative site of RFLP. Hence, the inventors took notice of the region capable of yielding a comparatively long PCR product (2.7 kb) and made an attempt to create a dCAPS marker.

[0147] Specifically, upon comparing the base sequences of the genomic PCR products of said region using Asominori and Koshihikari (both japonica rice) and Kasalath and IR24 (both indica rice), the inventors found six sites of polymorphism between japonica and indica. One of these six sites was used to create a dCAPS marker. To this end, with SEQ ID NO:5 and SEQ ID NO:6 used as primers, PCR was performed over 35 cycles, each cycle consisting of 94°C x 30 sec, 58°C x 30 sec and 72°C x 30 sec. The PCR product was treated with MwoI and analyzed by electrophoresis on 3% MetaPhor™ agarose gel. In Asominori, cleavage occurred at two sites to give three observable bands of about 25 bp, 50 bp and 79 bp, but in IR24 cleavage occurred at one site to give two observable bands of about 50 bp and 107 bp. Mapping demonstrated the conversion of RFLP marker C1361 to a PCR marker, which was named C1361 MwoI (SEQ ID NO:20) in the present invention.

(7) Conversion of G2155 to PCR marker

[0148] Primers specific to both ends of the marker probe were designed and PCR was performed with the total DNA of Asominori, Koshihikari, IR24 or IL216 (a line produced by introducing Rf-1 gene into Koshihikari by back crossing; its genotype was Rf-1/Rf-1) being used as a template. Purification of the PCR product and searching for a mutation that would be useful for providing restriction fragment polymorphisms were performed by the methods already described above.

[0149] Specifically, as a result of comparing the base sequences of corresponding regions of the varieties under test, mutations were found at three sites between the variety/line (IR24 and IL216) having Rf-1 gene and the variety (Asominori and Koshihikari) not having Rf-1 gene. One of the three sites was utilized to create a dCAPS marker. To this end, SEQ ID NO:7 and SEQ ID NO:8 were used as primers to perform PCR over 35 cycles, each cycle consisting of 94°C x 30 sec, 58°C x 30 sec and 72°C x 30 sec. The PCR product was treated with MwoI and analyzed by electrophoresis on 3% MetaPhor™ agarose gel. In Asominori, cleavage occurred at one site to give two observable bands of about 25 bp and 105 bp, but in IR24, cleavage occurred at two sites to give three observable bands of about 25 bp, 27 bp and 78 bp. Mapping demonstrated the conversion of RFLP marker G2155 to a PCR marker, which was named G2155 MwoI (SEQ ID NO:21) in the present invention.

(8) Conversion of G291 to PCR marker

[0150] Primers specific to internal sequences of the marker probe were designed and used in various combinations to perform PCR to find a primer combination that could yield an amplification product of the expected size. Using the selected primer combination, the inventors performed PCR with the total DNA of Asominori, Koshihikari, IR24 and IL216 used as a template. Purification of the PCR product and searching for a mutation that could be utilized in providing restriction fragment polymorphisms were performed by the methods already described above.

[0151] Specifically, using the primers designed to be specific for the marker probe sequence, the inventors performed genomic PCR of each variety under test and compared the base sequences of the products. As a result, mutations were found at four sites between the variety/line having Rf-1 gene (IR24 and IL216) and the variety (Asominori and Koshihikari) not having Rf-1 gene. One of the four sites was used to create a dCAPS marker. To this end, SEQ ID NO:9 and SEQ ID NO:10 were used as primers to perform PCR over 35 cycles, each cycle consisting of 94°C x 30 sec, 58°C x 30 sec and 72°C x 30 sec. The PCR product was treated with MspI and analyzed by electrophoresis on 3% MetaPhor™ agarose gel. In the varieties/lines having Rf-1 gene, cleavage occurred at two sites to give three observable bands of about 25 bp, 49 bp and 55 bp, but in the varieties not having Rf-1 gene, cleavage occurred at one site to give two observable bands of about 25 bp and 104 bp. Mapping demonstrated the conversion of RFLP marker G291 to a PCR marker, which was named G291 MspI (SEQ ID NO:22) in the present invention.

(9) Conversion of R2303 to PCR marker

[0152] Primers specific to internal sequences of the marker probe were designed and PCR was performed with the total DNA of Asominori (japonica rice) and IR24 and Kasalath (indica rice) used as a template. Purification of the PCR product and searching for a mutation that could be used for providing restriction fragment polymorphisms were performed by the methods already described above.

[0153] As a result of comparing the base sequences of corresponding regions of the varieties under test, a mutation was found between japonica rice and indica rice. Since the mutation occurred at the BstI recognition site, the site was directly used to create a CAPS marker. To this end, SEQ ID NO:11 and SEQ ID NO:12 were used as primers and PCR was performed over 30 cycles, each cycle consisting of 94°C x 1 min, 58°C x 1 min and 72°C x 2 min. The PCR product was treated with BstI and analyzed by electrophoresis on 2% agarose gel. In japonica rice, cleavage occurred at one site to give two observable bands of about 238 bp and 1334 bp, but in indica rice, cleavage occurred at two sites to give three observable bands of about 238 bp, 655 bp and 679 bp. Mapping demonstrated the conversion of RFLP marker R2303 to a PCR marker, which was named R2303 BstI (SEQ ID NO:23) in the present invention.

(10) Converting S10019 to PCR marker

[0154] S10019 was converted to a PCR marker in accordance with the method (9) of converting R2303 to a PCR marker.

[0155] Specifically, as a result of comparing the base sequences of corresponding regions of the varieties under test, a mutation was found between japonica rice and indica rice. Since the mutation occurred at the BstUI recognition site, the site was directly used to create a CAPS marker. To this end, SEQ ID NO:13 and SEQ ID NO:14 were used as primers and PCR was performed over 30 cycles, each cycle consisting of 94°C x 1 min, 58°C x 1 min and 72°C x 1

min. The PCR product was treated with BstUI and analyzed by electrophoresis on 2% agarose gel. In japonica rice, cleavage occurred at one site to give two observable bands of about 130 bp and 462 bp, but in indica rice, cleavage occurred at two sites to give three observable bands of about 130 bp, 218 bp and 244 bp. Mapping demonstrated the conversion of RFLP marker S10019 to a PCR marker, which was named S10019 BstUI (SEQ ID NO:24) in the present invention.

(11) Conversion of S10602 to PCR marker

[0156] S10602 was converted to a PCR marker in accordance with the method (9) of converting R2303 to a PCR marker.

[0157] Specifically, as a result of comparing the base sequences of corresponding regions of the varieties under test, a mutation was found between japonica rice and indica rice. The mutation was used to create a CAPS marker. To this end, SEQ ID NO:15 and SEQ ID NO:16 were used as primers and PCR was performed over 33 cycles, each cycle consisting of 94°C x 1 min, 58°C x 1 min and 72°C x 1 min. The PCR product was treated with KpnI and analyzed by electrophoresis on 2% agarose gel. In japonica rice, cleavage occurred at one site to give two observable bands of about 117 bp and 607 bp, but in indica rice, no cleavage occurred, giving only an observable band of 724 bp. Mapping demonstrated the conversion of RFLP marker S10602 to a PCR marker, which was named S10602 KpnI (SEQ ID NO:25) in the present invention.

(12) Conversion of S12564 to PCR marker

[0158] S12564 was converted to a PCR marker in accordance with the method of converting R2303 to a PCR marker.

[0159] Specifically, as a result of comparing the base sequences of corresponding regions of the varieties under test, a mutation was found between japonica rice and indica rice. The mutation was used to create a dCAPS marker. To this end, SEQ ID NO:17 and SEQ ID NO:18 were used as primers and PCR was performed over 35 cycles, each cycle consisting of 94°C x 30 sec, 58°C x 30 sec and 72°C x 30 sec. The PCR product was treated with Tsp509I and analyzed by electrophoresis on 3% MetaPhor™ agarose gel. In japonica rice, cleavage occurred at two sites to give three observable bands of 26 bp, 41 bp and 91 bp, but in indica rice, cleavage occurred at one site to give two observable bands of 41bp and 117 bp. Mapping demonstrated the conversion of RFLP marker S12564 to a PCR marker, which was named S12564 Tsp509I (SEQ ID NO:26) in the present invention.

Reference example 2: Mapping of Rf-1 Gene Locus

[0160] DNA was extracted from 1042 seedlings of the F1 population produced by pollinating MS Koshihikari with MS-FR Koshihikari, and the DNA extract was used in the analysis. MS Koshihikari (generation: BC10F1) was created by replacing the cytoplasm of Koshihikari with BT type male sterility cytoplasm. MS-FR Koshihikari was a line created by introducing Rf-1 gene from IR8 (supplied from National Institute of Agrobiological Sciences) into MS Koshihikari (the locus of Rf-1 gene being heterozygous).

[0161] First, each individual was investigated for the genotype at two marker loci R1877 EcoRI and G2155 MwoI described in Reference example 1 that would presumably be located on opposite sides of the locus of Rf-1 gene. Japonica type homozygotes with respect to either locus R1877 EcoRI or G2155 MwoI were regarded as recombinants between these two marker loci. Then, each of such recombinants was investigated for the genotypes of G291 MspI, R2303 BstI, S12564 Tsp 509I, C1361 MwoI, S10019 BstUI, G4003 HindIII and S10602 KpnI loci, and the positions of recombination were identified.

[0162] The genotype investigation with respect to R1877 EcoRI and G2155 MwoI loci revealed that 46 individuals were recombinants around the locus of Rf-1 gene. Genotypes of the marker loci around the locus of Rf-1 gene were investigated and the results are shown in Table 3.

**Table 3. Genotypes of Marker Loci in Recombinant
Individuals Around Rf-1 Locus**

Locus	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
R1877 EcoRI	J	J	J	J	J	J	J	J	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
G291 MspI	H	J	J	J	J	J	J	J	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
R2303 BstII	H	H	J	J	J	J	J	J	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
S12564 Tsp509I	H	H	H	H	H	H	H	J	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
C1361 MwoI	H	H	H	H	H	H	H	H	J	J	H	H	H	H	H	H	H	H	H	H	H	H	H
S10018 BstUI	H	H	H	H	H	H	H	H	J	J	J	J	J	J	J	J	H	H	H	H	H	H	H
G4003 HindIII	H	H	H	H	H	H	H	H	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J
S10602 KpnI	H	H	H	H	H	H	H	H	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J
G2155 MwoI	H	H	H	H	H	H	H	H	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J
	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
	J	J	J	J	J	J	J	J	J	H	H	H	H	H	H	H	H	H	H	H	H	H	H
	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	H	H
	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J	J

J: Homozygous Koshihikari type

H: Heterozygous Koshihikari type/ MS-FR Koshihikari type

[0163] As shown in Table 3, recombinant 8 homozygous for japonica at the S12564 Tsp509I marker locus and recombinants 9 and 10 homozygous for japonica at the C1361 MwoI marker locus were obtained. As all of these recombinants restored fertility, the former was regarded as a recombinant between the Rf-1 and S12564 Tsp509I loci while the latter were regarded as recombinants between the Rf-1 and C1361 MwoI loci, showing that the Rf-1 gene is located between the S12564 Tsp509I and C1361 MwoI loci. Based on the report that only pollens carrying the Rf-1 gene have fertility in individuals having the BT type male sterile cytoplasm in the cross above (C.Shinjo, JAPAN.J.GENETICS Vol.44, No.3:149-156(1969)), the Rf-1 gene locus could be located on a detailed linkage map (Fig. 4).

Example 1: Acquisition of recombinant individuals proximal to the Rf-1 locus

(Materials and Methods)

[0164] DNA was extracted from each of 4103 individuals of BC10F1 population produced by pollinating MS Koshihikari (generation: BC10F1) with MS-FR Koshihikari (generation: BC9F1, heterozygous at the Rf-1 locus), and genotyped at the S12564 Tsp509I and C1361 MwoI loci in the same manner as described in Reference example 2 above. Individuals having a genotype homozygous for Koshihikari at the S12564 Tsp509I locus were regarded as those generated by recombination between the Rf-1 and S12564 Tsp509I loci, while individuals having a genotype homozygous for Koshihikari at the C1361 MwoI locus were regarded as those generated by recombination between the Rf-1 and C1361 MwoI loci.

(Results and Discussion)

[0165] A survey of 4103 individuals revealed one recombinant individual between the Rf-1 and S12564 Tsp509I loci and 6 recombinant individuals between the Rf-1 and C1361 MwoI loci. The previous survey of 1042 individuals obtained by crossing in Reference example 2 above had already revealed one recombinant individual between the Rf-1 and S12564 Tsp509I loci and 2 recombinant individuals between the Rf-1 and C1361 MwoI loci as shown in Table 3.

[0166] Thus, a total of 2 recombinant individuals between the Rf-1 and S12564 Tsp509I loci and 8 recombinant individuals between the Rf-1 and C1361 MwoI loci were able to be obtained from 5145 individuals. These 10 individuals

were tested by high-precision segregation analysis in the examples below.

Example 2 chromosomal walking

(1) First chromosomal walking

(Materials and Methods)

[0167] A genomic library was constructed from the genomic DNA of Asominori japonica (not carrying Rf-1) using Lambda DASH II vector as described in Reference example 1 and tested by chromosomal walking.

[0168] PCR was routinely performed using total DNA of Asominori as a template in combination with the following primer pair:

5'-atcaggagccttcaaattgggaac-3' (SEQ ID NO: 29)

and

5'-ctcgcaaattgottaattttgacc-3' (SEQ ID NO: 30)

designed for a partial base sequence (Accession No. D47284) of RFLP probe S12564. The resulting amplification products of about 1200 bp were electrophoresed on an agarose gel and then purified by QIAEXII (QIAGEN). The purified DNA was labeled with a rediprime DNA labelling system (Amersham Pharmacia) to give a library screening probe (probe A, Fig. 1).

[0169] The library was routinely screened after plaques were blotted onto Hybond-N⁺ (Amersham Pharmacia). Single plaques were separated, after which phage DNA was purified by the plate lysate method using Lambda Midi kit (QIAGEN).

(Results and Discussion)

[0170] The results of terminal base sequence analysis and restriction enzyme fragment length analysis showed that two (WSA1 and WSA3) of 4 clones obtained by screening were in a relative position as shown in Fig. 1. The Asominori genomic base sequences corresponding to WSA1 and WSA3 were determined by primer walking (DNA Sequencer 377, ABI).

(2) Second chromosomal walking

(Materials and Methods)

[0171] In addition to the Asominori genomic library described above, an IR24 genomic library was similarly constructed from the genomic DNA of an indica variety IR24 (carrying Rf-1) and tested by chromosomal walking.

[0172] PCR was routinely performed using DNA of WSA3 as a template in combination with the following primer pair:

5'-tgaaggagttatgggtgcgtgacg-3' (SEQ ID NO: 31)

and

5'-ttgccgagcacacttgccatgtgc-3' (SEQ ID NO: 32)

designed for the Asominori genomic base sequence determined in (1). The resulting amplification products of 524 bp were purified and labeled by the method described above to give a library screening probe (probe E, Fig. 1).

[0173] Library screening and phage DNA purification were performed by the method described above.

(Results and Discussion)

[0174] The results of terminal base sequence analysis and restriction enzyme fragment length analysis showed that one (WSE8) of 15 clones obtained by screening of the Asominori genomic library was in a relative position as shown in Fig. 1. The Asominori genomic base sequence corresponding to WSE8 was determined by primer walking.

[0175] The results of terminal base sequence analysis and restriction enzyme fragment length analysis showed that two (XSE1 and XSE7) of 7 clones obtained by screening of the IR24 genomic library were in a relative position as shown in Fig. 1. The IR24 genomic base sequences corresponding to XSE1 and XSE7 were determined by primer walking.

(3) Third chromosomal walking

(Materials and Methods)

[0176] The Asominori genomic library and IR24 genomic library described above were tested by chromosomal walking.

[0177] PCR was routinely performed using DNA of WSE8 as a template in combination with the following primer pair:

5' -gcgacgcaatggacatagtgtcc-3' (SEQ ID NO: 33)

and

5' -ttacctgccaagcaatatccatcg-3' (SEQ ID NO: 34)

designed for the Asominori genomic base sequence determined in (2). The resulting amplification products of 1159 bp were purified and labeled by the method described above to give a library screening probe (probe F, Fig. 1).

[0178] Library screening and phage DNA purification were performed by the method described above.

(Results and Discussion)

[0179] The results of terminal base sequence analysis and restriction enzyme fragment length analysis showed that two (WSF5 and WSF7) of 8 clones obtained by screening of the Asominori genomic library were in a relative position as shown in Fig. 1. The Asominori genomic base sequences corresponding to WSF5 and WSF7 were determined by primer walking.

[0180] The results of terminal base sequence analysis and restriction enzyme fragment length analysis showed that two (XSF4 and XSF20) of 13 clones obtained by screening of the IR24 genomic library were in a relative position as shown in Fig. 1. The IR24 genomic base sequences corresponding to XSF4 and XSF20 were determined by primer walking.

(4) Fourth chromosomal walking

(Materials and Methods)

[0181] The Asominori genomic library and IR24 genomic library described above were tested by chromosomal walking.

[0182] PCR was routinely performed using DNA of WSF7 as a template in combination with the following primer pair:

5' -aaggcatactcagtgaggaggcaag-3' (SEQ ID NO: 35)

and

5' -ttaacctgaccgcaagcacctgtc-3' (SEQ ID NO: 36)

designed for the Asominori genomic base sequence determined in (3). The resulting amplification products of

456 bp were purified and labeled by the method described above to give a library screening probe (probe G, Fig. 1).

[0183] Library screening and phage DNA purification were performed by the method described above.

(Results and Discussion)

[0184] The results of terminal base sequence analysis and restriction enzyme fragment length analysis showed that two (WSG2 and WSG6) of 6 clones obtained by screening of the Asominori genomic library were in a relative position as shown in Fig. 1. The Asominori genomic base sequences corresponding to WSG2 and WSG6 were determined by primer walking.

[0185] The results of terminal base sequence analysis and restriction enzyme fragment length analysis showed that three (XSG8, XSG16 and XSG22) of 14 clones obtained by screening of the IR24 genomic library were in a relative position as shown in Fig. 1. The IR24 genomic base sequences corresponding to XSG8, XSG16 and XSG22 were determined by primer walking.

(5) Fifth chromosomal walking

(Materials and Methods)

[0186] The IR24 genomic library described above was tested by chromosomal walking.

[0187] We perused the public website of TIGR (The Institute for Genomic Research) and found that a BAC (Bacterial Artificial Chromosome) clone (Accession No. AC068923) containing RFLP marker S12564 had been deposited with a public database (GenBank). This BAC clone contains the genomic DNA of *Nipponbare japonica* and it was shown from base sequence comparison to completely include the contig regions of Asominori and IR24 prepared in (1)-(4) (Fig. 2).

[0188] Thus, PCR was routinely performed using total DNA of IR24 as a template in combination with the following primer pair:

5' - tggatggactatgtgggggtcagtc - 3' (SEQ ID NO: 37)

and

5' - agtgggaagtggagagagtagggag - 3' (SEQ ID NO: 38)

designed to amplify a part of this BAC clone. The resulting amplification products of about 600 bp were purified and labeled by the method described above to give a library screening probe (probe H, Fig. 1).

[0189] Library screening and phage DNA purification were performed by the method described above.

(Results and Discussion)

[0190] The results of terminal base sequence analysis and restriction enzyme fragment length analysis showed that one (XSH18) of 15 clones obtained by screening of the IR24 genomic library was in a relative position as shown in Fig. 1. The IR24 genomic base sequence corresponding to XSH18 was determined by primer walking.

Example 3: High precision segregation analysis

(1) Development of PCR marker P4497 Mbol

[0191] Comparison between the genomic base sequence corresponding to the IR24 contig (SEQ ID NO: 27) and the genomic base sequence corresponding to the Asominori contig (SEQ ID NO: 28) determined in Example 2 revealed that the 1239th base, of SEQ ID NO: 27 is A while the 12631st base of SEQ ID NO: 28 corresponding to said position is G.

[0192] For detecting this change, fragments of about 730 bp are first amplified by PCR from a region surrounding said position using the following primer pair: P4497 Mbol F:

5' -ccctccaacacataaatggttgag-3' (SEQ ID NO: 39)

5 (corresponding to bases 853-876 of SEQ ID NO: 27)
(corresponding to bases 12247-12270 of SEQ ID NO: 28) and
P4497 Mbol R:

10 **5' -tttctgccaggaaactgtagatg-3' (SEQ ID NO: 40)**

(corresponding to bases 1583-1560 of SEQ ID NO: 27)
(corresponding to bases 12975-12952 of SEQ ID NO: 28).

15 The amplification products can be visualized by electrophoresis on an agarose gel after treatment with Mbol. Thus, the change can be detected as a difference in mobility in the agarose gel due to the difference in the length of DNA after Mbol treatment because the amplification products from Asominori DNA having an Mbol recognition sequence (GATC) are cleaved with Mbol while the amplification products from IR24 DNA are not cleaved with Mbol for the lack of the Mbol recognition sequence.

20 (2) Development of PCR marker P9493 BslI

[0193] Comparison between the genomic base sequence corresponding to the IR24 contig (SEQ ID NO: 27) and the genomic base sequence corresponding to the Asominori contig (SEQ ID NO: 28) determined in Example 2 revealed that the 6227th base of SEQ ID NO: 27 is A while the 17627th base of SEQ ID NO: 28 corresponding to said position is C.

25 [0194] For detecting this change, fragments of 126 bp are first amplified by PCR from a region surrounding said position using the following primer pair:
P9493 BslI F:

30 **5' -gcatctttatacgcatactatgcg-3' (SEQ ID NO: 41)**

(corresponding to bases 6129-6152 of SEQ ID NO: 27)
(corresponding to bases 17529-17552 of SEQ ID NO: 28) and

35 P9493 BslI R:

5' -aaagtctttgttccttcaccaagg-3' (SEQ ID NO: 42)

40 (corresponding to bases 6254-6231 of SEQ ID NO: 27)
(corresponding to bases 17654-17631 of SEQ ID NO: 28).

The amplification products can be visualized by electrophoresis on an agarose gel after treatment with BslI. Thus, the change can be detected as a difference in mobility in the agarose gel due to the difference in the length of DNA after BslI treatment because the amplification products from Asominori DNA having a BslI recognition sequence (CC-NNNNNNNGG) are cleaved with BslI while the amplification products from IR24 DNA are not cleaved with BslI for the lack of the BslI recognition sequence.

45 [0195] This marker was developed by applying the dCAPS method (Michaels and Amasino 1998, Neff et al., 1998). Specifically, g is substituted for a at the base 6236 of SEQ ID NO: 27 and the base 17636 of SEQ ID NO: 28 by the use of P9493 BslI R primer described above. Thus, the fragments from Asominori DNA come to have a sequence of CcttctctGG at 17626-17636 of SEQ ID NO: 28 so that they are cleaved with BslI.

(3) Development of PCR marker P23945 Mbol

55 [0196] Comparison between the genomic base sequence corresponding to the IR24 contig (SEQ ID NO: 27) and the genomic base sequence corresponding to the Asominori contig (SEQ ID NO: 28) determined in Example 2 revealed that the 20680th base of SEQ ID NO: 27 is G while the 32079th base of SEQ ID NO: 28 corresponding to said position is A.

[0197] For detecting this change, fragments of 260 bp are first amplified by PCR from a region surrounding said

position using the following primer pair: P23945 MboI F:

5' -gaggatttatcaaaacaggatggacg-3' (SEQ ID NO: 43)

(corresponding to bases 20519-20544 of SEQ ID NO: 27)

(corresponding to bases 31918-31943 of SEQ ID NO: 28) and
P23945 MboI R:

5' -tgggcggcagcagtggaggataga-3' (SEQ ID NO: 44)

(corresponding to bases 20778-20755 of SEQ ID NO: 27)

(corresponding to bases 32177-32154 of SEQ ID NO: 28).

The amplification products can be visualized by electrophoresis on an agarose gel after treatment with MboI. Thus, the change can be detected as a difference in mobility in the agarose gel due to the difference in the length of DNA after MboI treatment because the amplification products from IR24 DNA having an MboI recognition sequence (GATC) are cleaved with MboI while the amplification products from Asominori DNA are not cleaved with MboI for the lack of the MboI recognition sequence.

(4) Development of PCR marker P41030 TaqI

[0198] Comparison between the genomic base sequence corresponding to the IR24 contig (SEQ ID NO: 27) and the genomic base sequence corresponding to the Asominori contig (SEQ ID NO: 28) determined in Example 2 revealed that the 45461st base of SEQ ID NO: 27 is A while the 49164th base of SEQ ID NO: 28 corresponding to said position is G.

[0199] For detecting this change, fragments of 280 bp are first amplified by PCR from a region surrounding said position using the following primer pair:

P41030 TaqI F:

5' -aagaaggagggttatagaatctg-3' (SEQ ID NO: 45)

(corresponding to bases 45369-45392 of SEQ ID NO: 27)

(corresponding to bases 49072-49095 of SEQ ID NO: 28) and
P41030 TaqI R:

5' -atatcaggactaacaccactgctc-3' (SEQ ID NO: 46)

(corresponding to bases 45648-45625 of SEQ ID NO: 27)

(corresponding to bases 49351-49328 of SEQ ID NO: 28).

The amplification products can be visualized by electrophoresis on an agarose gel after treatment with TaqI. Thus, the change can be detected as a difference in mobility in the agarose gel due to the difference in the length of DNA after TaqI treatment because the amplification products from Asominori DNA having a TaqI recognition sequence (TCGA) are cleaved with TaqI while the amplification products from IR24 DNA are not cleaved with TaqI for the lack of the TaqI recognition sequence.

(5) Development of PCR marker P45177 BstUI

[0200] Comparison between the genomic base sequence corresponding to the IR24 contig (SEQ ID NO: 27) and the genomic base sequence corresponding to the Asominori contig (SEQ ID NO: 28) determined in Example 2 revealed that the 49609th base of SEQ ID NO: 27 is A while the 53311st base of SEQ ID NO: 28 corresponding to said position is G.

[0201] For detecting this change, fragments of 812 bp are first amplified by PCR from a region surrounding said

position using the following primer pair:

P45177 BstUI F:

5'-acgagtagtagcgatcttccagcg-3' (SEQ ID NO: 47)

(corresponding to bases 49355-49378 of SEQ ID NO: 27)

(corresponding to bases 53057-53080 of SEQ ID NO: 28) and

P45177 BstUI R:

5'-cagcgtgaaactaaaaacggaggc-3' (SEQ ID NO: 48)

(corresponding to bases 50166-50143 of SEQ ID NO: 27)

(corresponding to bases 53868-53845 of SEQ ID NO: 28).

The amplification products can be visualized by electrophoresis on an agarose gel after treatment with BstUI. Thus, the change can be detected as a difference in mobility in the agarose gel due to the difference in the length of DNA after BstUI treatment because the amplification products from IR24 DNA having a BstUI recognition sequence (CGCG) at two positions are cleaved into 3 fragments with BstUI while the amplification products from Asominori DNA having the BstUI recognition sequence at three positions are cleaved with BstUI into four fragments.

(6) Development of PCR marker B60304 MspI

[0202] Comparison between the genomic base sequence corresponding to the IR24 contig (SEQ ID NO: 27) determined in Example 2 and the base sequence of the BAC clone described above (Accession No. AC068923) revealed that the 56368th base of SEQ ID NO: 27 is T while the base of AC068923 corresponding to said position is C.

[0203] For detecting this change, fragments of about 330 bp are first amplified by PCR from a region surrounding said position using the following primer pair:

B60304 MspI F:

5'-atcccacatcatcataatccgacc-3' (SEQ ID NO: 49)

(corresponding to bases 56149-56172 of SEQ ID NO: 27) and

B60304 MspI R:

5'-agcttctcccttggatacgggtggcg-3' (SEQ ID NO: 50)

(corresponding to bases 56479-56455 of SEQ ID NO: 27).

The amplification products can be visualized by electrophoresis on an agarose gel after treatment with MspI. Thus, the change can be detected as a difference in mobility in the agarose gel due to the difference in the length of DNA after MspI treatment because the amplification products from Nipponbare DNA having an MspI recognition sequence (CCGG) are cleaved with MspI while the amplification products from IR24 DNA are not cleaved with MspI for the lack of the MspI recognition sequence.

[0204] This marker was developed by applying the dCAPS method. Specifically, t is substituted for g at base 56463 of SEQ ID NO: 27 by the use of B60304 MspI R primer. As a result, the MspI recognition sequence of bases 56460-56463 of SEQ ID NO: 27 changes from CCGG into ccgt so that the fragments from SEQ ID NO: 27 become unable to be cleaved with MspI. Thus, the fragments from IR24 have no MspI recognition sequence, while DNA from Nipponbare has the MspI recognition sequence at one position in a region corresponding to bases 56367-56370 of SEQ ID NO: 27.

(7) Development of PCR marker B59066 BsaJI

[0205] Comparison between the genomic base sequence corresponding to the IR24 contig (SEQ ID NO: 27) determined in Example 2 and the base sequence of the BAC clone described above (Accession No. AC068923) revealed

that the 57629th base of SEQ ID NO: 27 is C while the base of AC068923 corresponding to said position is CC.
[0206] For detecting this change, fragments of about 420 bp are first amplified by PCR from a region surrounding said position using the following primer pair:
 B59066 BsaJI F:

5'-atttggttggttagttgcggctgag-3' (SEQ ID NO: 51)

(corresponding to bases 57563-57586 of SEQ ID NO: 27) and
 B59066 BsaJI R:

5'-gcccaaactcaaaaggagagaacc-3' (SEQ ID NO: 52)

(corresponding to bases 57983-57960 of SEQ ID NO: 27).
 The amplification products can be visualized by electrophoresis on an agarose gel after treatment with BsaJI. Thus, the change can be detected as a difference in mobility in the agarose gel due to the difference in the length of DNA after BsaJI treatment because the amplification products from Nipponbare DNA having a BsaJI recognition sequence (CCNNGG) are cleaved with BsaJI while the amplification products from IR24 DNA are not cleaved with BsaJI for the lack of the BsaJI recognition sequence.

(8) Development of PCR marker B56691 XbaI

[0207] Comparison between the genomic base sequence corresponding to the IR24 contig (SEQ ID NO: 27) determined in Example 2 and the base sequence of the BAC clone described above (Accession No. AC068923) revealed that the 66267th base of SEQ ID NO: 27 is G while the base of AC068923 corresponding to said position is C.

[0208] For detecting this change, fragments of about 670 bp are first amplified by PCR from a region surrounding said position using the following primer pair:
 B56691 XbaI F:

5'-cctcaagtctccctaaagccact-3' (SEQ ID NO: 53)

(corresponding to bases 66129-66152 of SEQ ID NO: 27) and
 B56691 XbaI R:

5'-gctctactgctgataaacggtgag-3' (SEQ ID NO: 54)

(corresponding to bases 66799-66776 of SEQ ID NO: 27).
 The amplification products can be visualized by electrophoresis on an agarose gel after treatment with XbaI. Thus, the change can be detected as a difference in mobility in the agarose gel due to the difference in the length of DNA after XbaI treatment because the amplification products from Nipponbare DNA having an XbaI recognition sequence (TCTAGA) are cleaved with XbaI while the amplification products from IR24 DNA are not cleaved with XbaI for the lack of the XbaI recognition sequence.

(9) Development of PCR marker B53627 BstZ17I

[0209] Comparison between the genomic base sequence corresponding to the IR24 contig (SEQ ID NO: 27) determined in Example 2 and the base sequence of the BAC clone described above (Accession No. AC068923) revealed that the 69331st base of SEQ ID NO: 27 is T while the base of AC068923 corresponding to said position is C.

[0210] For detecting this change, fragments of about 620 bp are first amplified by PCR from a region surrounding said position using the following primer pair:
 B53627 BstZ17I F:

5'-tggatggactatgtgggggtcagtc-3' (SEQ ID NO: 55)

(corresponding to bases 68965-68988 of SEQ ID NO: 27) and
B53627 BstZ17I R:

5'-agtggaagtggagagagtagggag-3' (SEQ ID NO: 56)

(corresponding to bases 69582-69559 of SEQ ID NO: 27).

The amplification products can be visualized by electrophoresis on an agarose gel after treatment with BstZ17I. Thus, the change can be detected as a difference in mobility in the agarose gel due to the difference in the length of DNA after BstZ17I treatment because the amplification products from IR24 DNA having a BstZ17I recognition sequence (GTATAC) are cleaved with BstZ17I while the amplification products from Nipponbare DNA are not cleaved with BstZ17I for the lack of the BstZ17I recognition sequence.

(10) Development of PCR marker B40936 MseI

[0211] Development of all the following PCR markers (10)-(12) relates to a study of the base sequences corresponding to further downstream regions (3') of base 76363 at the 3' end of SEQ ID NO: 27.

[0212] The following primer pair was designed for the base sequence of the BAC clone described above (Accession No. AC068923):

5'-tacgacgccatttcactccattgc-3' (SEQ ID NO: 57)

and

5'-catttctctatgggcgttgctctg-3' (SEQ ID NO: 58).

PCR was routinely performed using this primer pair in combination with total DNAs of MS-FR Koshihikari (genotype of the Rf-1 locus: Rf-1 Rf-1) and Koshihikari as templates. The resulting amplification products of about 1300 bp were electrophoresed on an agarose gel and then purified by QIAEXII (QIAGEN). Analysis of the base sequence of the purified DNA by a DNA sequencer 377 (ABI) showed several polymorphisms.

[0213] One of them can be detected by PCR amplification of a region surrounding said position using the following primer pair:
B40936 MseI F:

5'-acctgtaggtatggcaccttcaacac-3' (SEQ ID NO: 59)

and
B40936 MseI R:

5'-ccaaggaacgaagttcaaatgtatgg-3' (SEQ ID NO: 60).

The amplification products can be visualized by electrophoresis on an agarose gel after treatment with MseI. Thus, the change can be detected as a difference in mobility in the agarose gel due to the difference in the length of DNA after MseI treatment because the amplification products from MS-FR Koshihikari (Rf-1 Rf-1) DNA having an MseI recognition sequence (TTAA) are cleaved with MseI while the amplification products from Koshihikari DNA are not cleaved with MseI for the lack of the MseI recognition sequence.

[0214] This marker was developed by applying the dCAPS method.

(11) Development of PCR marker B19839 MwoI

[0215] The following primer pair was designed for the base sequence of the BAC clone described above (Accession No. AC068923):

5'-tgatgtgtttgggcatccctttcg-3' (SEQ ID NO: 61)

and

5'-gagataggggacgacagacacgac-3' (SEQ ID NO: 62).

PCR was routinely performed using this primer pair in combination with total DNAs of MS-FR Koshihikari (genotype of the Rf-1 locus: Rf-1 Rf-1) and Koshihikari as templates. The resulting amplification products of about 1200 bp were electrophoresed on an agarose gel and then purified by QIAEXII (QIAGEN). Analysis of the base sequence of the purified DNA by a DNA sequencer 377 (ABI) showed several polymorphisms.

[0216] One of them can be detected by PCR amplification of a region surrounding said position using the following primer pair:

B19839 MwoI F:

5'-tcctatggctgtttagaaactgcaca-3' (SEQ ID NO: 63)

and

B19839 MwoI R:

5'-caagttcaaacataactggcggtg-3' (SEQ ID NO: 64).

The amplification products can be visualized by electrophoresis on an agarose gel after treatment with MwoI. Thus, the change can be detected as a difference in mobility in the agarose gel due to the difference in the length of DNA after MwoI treatment because the amplification products from Koshihikari DNA having an MwoI recognition sequence (GCNNNNNNNGC) are cleaved with MwoI while the amplification products from MS-FR Koshihikari (Rf-1 Rf-1) DNA are not cleaved with MwoI for the lack of the MwoI recognition sequence.

[0217] This marker was developed by applying the dCAPS method.

(12) Development of PCR marker B2387 BfaI

[0218] The following primer pair was designed for the base sequence of the BAC clone described above (Accession No. AC068923):

5'-cactgtcctgtaagtgtgctgtgc-3' (SEQ ID NO: 65)

and

5'-caagcgtgtgataaaatgtgacgc-3' (SEQ ID NO: 66).

PCR was routinely performed using this primer pair in combination with total DNAs of MS-FR Koshihikari (genotype of the Rf-1 locus: Rf-1 Rf-1) and Koshihikari as templates. The resulting amplification products of about 1300 bp were electrophoresed on an agarose gel and then purified by QIAEXII (QIAGEN). Analysis of the base sequence of the purified DNA by a DNA sequencer 377 (ABI) showed several polymorphisms.

[0219] One of them can be detected by PCR amplification of a region surrounding said position using the following primer pair:

B2387 BfaI F:

5'-tgcctactgccattactatgtgac-3' (SEQ ID NO: 67)

and
B2387 Bfal R:

5' -acatactaccgtaaattgggtctctg-3' (SEQ ID NO: 68).

The amplification products can be visualized by electrophoresis on an agarose gel after treatment with Bfal. Thus, the change can be detected as a difference in mobility in the agarose gel due to the difference in the length of DNA after Bfal treatment because the amplification products from Koshihikari DNA having an Bfal recognition sequence (CTAG) are cleaved with Bfal while the amplification products from MS-FR Koshihikari (Rf-1 Rf-1) DNA are not cleaved with Bfal for the lack of the Bfal recognition sequence.

(13) Segregation analysis

[0220] Two recombinants between the Rf-1 and S12564 Tsp509I loci (RS1 and RS2) and 8 recombinants between the Rf-1 and C1361 MwoI loci (RC1 to RC8) obtained in Example 1 were genotyped at the 12 DNA marker loci developed in (1) to (12) above. The results are shown in Table 4 along with the genotypes of each recombinant at the S12564 Tsp509I and C1361 MwoI loci.

Table 4.

Genotypes of recombinants proximal to the Rf-1 locus at various marker loci										
Locus	RS1	RS2	RC1	RC2	RC3	RC4	RC5	RC6	RC7	RC8
S12564 Tsp509I	J	J	H	H	H	H	H	H	H	H
P4497 MboI	J	J	H	H	H	H	H	H	H	H
P9493 BstII	H	H	H	H	H	H	H	H	H	H
P23945 MboI	H	H	H	H	H	H	H	H	H	H
P41030 TaqI	H	H	H	H	H	H	H	H	H	H
P45177 BstUI	H	H	H	H	H	H	H	H	H	H
B60304 MspI	H	H	H	H	H	H	H	H	H	H
B59066 BsaJI	H	H	H	H	H	H	H	H	H	H
B56691 XbaI	H	H	H	H	H	H	H	J	H	H
B53627 BstZ17I	H	H	H	H	H	H	H	J	H	H
B40936 MseI	H	H	H	H	H	H	H	J	H	H
B19839 MwoI	H	H	H	H	H	J	H	J	H	H
B2387 Bfal	H	H	H	H	H	J	R	J	H	J
C1361 MwoI	H	H	J	J	J	J	J	J	J	J
J: Homozygous for Koshihikari										
H: Heterozygous for Koshihikari /MS-FR Koshihikari										

[0221] Table 4 shows that all the recombinants have an indica-derived Rf-1 chromosomal region between P9493 BstII and 59066 BsaJI. This result showed that recombinant pollens having the chromosomal organization as shown in Fig. 3 have pollen fertility, i.e. the Rf-1 gene is functional in these pollens. This means that a sequence determining the presence of the function of the Rf-1 gene is included in the indica region common to these recombinant pollens, i.e. in a region from the P4497 MboI to B56691 XbaI loci (about 65 kb) as estimated at maximum.

[0222] However, there is a possibility that it is important for the expression of the genetic function of the Rf-1 gene that the Rf-1 gene is partially of the indica genotype, and that the genetic function may not be significantly changed whether the remaining regions are of the japonica or indica genotype. Therefore, it cannot be concluded that the common indica region above (bases 1239-66267 of SEQ ID NO: 27) completely contains the entire Rf-1 gene. However, it is thought that at least SEQ ID NO: 27 completely contains the entire Rf-1 gene for the following reasons:

- 1) the size of a gene is normally several kilobases, and rarely exceeds 10 kb;
- 2) the genomic base sequence of IR24 determined by the present invention (SEQ ID NO: 27) completely contains the common indica region above;
- 3) the 5' end of SEQ ID NO: 27 is located 1238 bp upstream of the 5' end of the common indica region above and

forms a part of another gene (S12564); and

4) the 3' end of SEQ ID NO: 27 is located 10096 bp downstream of the 3' end of the common indica region above.

Example 4: Complementation assay for a 9.7 kb fragment from XSE1

(Materials and Methods)

[0223] The λ phage clone XSE1 (Figs. 1 and 5) was completely digested with NotI and electrophoresed on an agarose gel. The separated 9.7 kb fragment (including bases 1-9657 of SEQ ID NO: 27) was purified by QIAEXII (QIAGEN).

[0224] On the other hand, an intermediate vector pSB200 having a hygromycin-resistant gene cassette was prepared on the basis of pSB11 (Komari et al., supra.). Specifically, a nopaline synthase terminator (Tnos) was first fused to a ubiquitin promoter and a ubiquitin intron (Pubi-ubil). A hygromycin-resistant gene (HYG(R)) was inserted between ubil and Tnos of the resulting Pubi-ubil-Tnos complex to give an assembly of Pubi-ubil-HYG(R)-Tnos. This assembly was fused to a HindIII/EcoRI fragment of pSB11 to give pKY205. Linker sites for adding restriction enzyme sites NotI, NspV, EcoRV, KpnI, SacI, EcoRI were inserted into the Hind III site upstream of Pubi of this pKY205 to give pSB200 having a hygromycin-resistant gene cassette.

[0225] After the plasmid vector pSB200 was completely digested with NotI, DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then dephosphorylated by CIAP (TAKARA). The reaction solution was electrophoresed on an agarose gel, and then a vector fragment was purified from the gel using QIAEXII (QIAGEN).

[0226] The two fragments prepared above, i.e. a 9.7 kb fragment from XSE1 and a vector fragment were subjected to a ligation reaction using DNA Ligation Kit Ver. 1 (TAKARA). After the reaction, DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in pure water (prepared by a Millipore system) and then mixed with *E. coli* DH5a cells, and the mixture was electroporated. After electroporation, the solution was cultured with shaking in LB medium (37°C, 1 hr) and then plated on an LB plate containing spectinomycin and warmed (37°C, 16 hr). Plasmids were isolated from 24 of the resulting colonies. Their restriction enzyme fragment length patterns and boundary base sequences were analyzed to select desired *E. coli* cells transformed with recombinant plasmids.

[0227] The *E. coli* cells selected above were used for triparental mating with the *Agrobacterium tumefaciens* strain LBA4404/pSB1 (Komari et al., 1996) and the helper *E. coli* strain HB101/pRK2013 (Ditta et al., 1980) according to the method of Ditta et al. (1980). Plasmids were isolated from 6 of the colonies formed on an AB plate containing spectinomycin and their restriction enzyme fragment length patterns were analyzed to select desired *Agrobacterium* cells.

[0228] The *Agrobacterium* cells selected above were used to transform MS Koshihikari (having BT cytoplasm and a nucleus gene substantially identical to Koshihikari) according to the method of Hiei et al. (1994). Necessary immature seeds of MS Koshihikari for transformation can be prepared by pollinating MS Koshihikari with Koshihikari.

[0229] Transformed plants were transferred to a greenhouse under long-day conditions after acclimation. 48 individuals grown to a stage suitable for transplantation were transplanted into 1/5000a Wagner pots (4 individuals/pot), and transferred into a greenhouse under short-day conditions 3-4 weeks after transplantation. About one month after heading, seed fertility was tested on standing plants.

(Results and Discussion)

[0230] All of the 48 transformed individuals were sterile. This indicates that the 9.7 kb insert fragment does not contain at least the full-length Rf-1 gene.

Example 5: Complementation assay for a 14.7 kb fragment from XSE7

(Materials and Methods)

[0231] The λ phage clone XSE7 (Figs. 1 and 5) was completely digested with EcoRI and then DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then blunted by DNA Blunting Kit (TAKARA). The reaction solution was electrophoresed on an agarose gel to separate a 14.7 kb fragment (including bases 2618-17261 of SEQ ID NO: 27), which was purified by QIAEXII (QIAGEN).

[0232] On the other hand, the plasmid vector pSB200 was completely digested with SacI and then DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then dephosphorylated by CIAP (TAKARA) and DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then blunted by DNA Blunting Kit (TAKARA). The reaction solution was electrophoresed on an agarose gel, and then a vector fragment was purified from the gel using QIAEXII (QIAGEN).

[0233] The two fragments prepared above, i.e. the 14.7 kb fragment from XSE7 and the vector fragment were sub-

jected to a ligation reaction using DNA Ligation Kit Ver. 1 (TAKARA). Subsequently, transformed plants were prepared and studied according to the method described in Example 4.

(Results and Discussion)

[0234] All of the 48 transformed individuals were sterile. This indicates that the 14.7 kb insert fragment does not contain at least the full-length Rf-1 gene.

Example 6: Complementation assay for a 21.3 kb fragment from XSF4

(Materials and Methods)

[0235] The λ phage clone XSF4 (Figs. 1 and 5) was partially digested with NotI and electrophoresed on an agarose gel. The separated 21.3 kb fragment (including bases 12478-33750 of SEQ ID NO: 27) was purified by QIAEXII (QIAGEN).

[0236] On the other hand, the plasmid vector pSB200 was completely digested with NotI and then DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then dephosphorylated by CIAP (TAKARA). The reaction solution was electrophoresed on an agarose gel, and then a vector fragment was purified from the gel using QIAEXII (QIAGEN).

[0237] The two fragments prepared above, i.e. the 21.3 kb fragment from XSF4 and the vector fragment were subjected to a ligation reaction using DNA Ligation Kit Ver. 1 (TAKARA). Subsequently, transformed plants were prepared and studied according to the method described in Example 4.

(Results and Discussion)

[0238] All of the 48 transformed individuals were sterile. This indicates that the 21.3 kb insert fragment does not contain at least the full-length Rf-1 gene.

Example 7: Complementation assay for a 13.2 kb fragment from XSF20

(Materials and Methods)

[0239] The λ phage clone XSF20 (Figs. 1 and 5) was completely digested with Sall and then DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then blunted by DNA Blunting Kit (TAKARA). The reaction solution was electrophoresed on an agarose gel to separate a 13.2 kb fragment (including bases 26809-40055 of SEQ ID NO: 27), which was purified by QIAEXII (QIAGEN).

[0240] On the other hand, the plasmid vector pSB200 was completely digested with EcoRV and then DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then dephosphorylated by CIAP (TAKARA). The reaction solution was electrophoresed on an agarose gel, and then a vector fragment was purified from the gel using QIAEXII (QIAGEN).

[0241] The two fragments prepared above, i.e. the 13.2 kb fragment from XSF20 and the vector fragment were subjected to a ligation reaction using DNA Ligation Kit Ver. 1 (TAKARA). Subsequently, transformed plants were prepared and studied according to the method described in Example 4.

(Results and Discussion)

[0242] All of the 44 transformed individuals were sterile. This indicates that the 13.2 kb insert fragment does not contain at least the full-length Rf-1 gene.

Example 8: Complementation assay for a 16.2 kb fragment from XSF18

(Materials and Methods)

[0243] The λ phage clone XSF18 is identical to XSF20 at the 5' and 3' ends (bases 20328 and 41921 of SEQ ID NO: 27, respectively), but lacks internal bases 33947-38591. Thus, it comprises bases 20328-33946 and 38592-41921 of SEQ ID NO: 27. This is because clone XSF18 was initially isolated but found to contain the above deletion during amplification after isolation, and therefore, the amplification step was freshly taken to isolate a complete clone designated XSF20.

[0244] The λ phage clone XSF18 (Fig. 5) was completely digested with NotI and electrophoresed on an agarose gel. The separated 16.2 kb fragment (including bases 21065-33946 and 38592-41921 of SEQ ID NO: 27) was purified by QIAEXII (QIAGEN).

[0245] On the other hand, the plasmid vector pSB200 was completely digested with NotI and then DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then dephosphorylated by CIAP (TAKARA). The reaction solution was electrophoresed on an agarose gel, and then a vector fragment was purified from the gel using QIAEXII (QIAGEN).

[0246] The two fragments prepared above, i.e. the 16.2 kb fragment from XSF18 and the vector fragment were subjected to a ligation reaction using DNA Ligation Kit Ver. 1 (TAKARA). Subsequently, transformed plants were prepared and studied according to the method described in Example 4.

(Results and Discussion)

[0247] All of the 48 transformed individuals were sterile (Fig. 6). This indicates that the 16.2 kb insert fragment does not contain at least the full-length Rf-1 gene.

Example 9: Complementation assay for a 12.6 kb fragment from XSG22

(Materials and Methods)

[0248] The λ phage clone XSG22 (Figs. 1 and 5) was partially digested with NotI and electrophoresed on an agarose gel. The separated 12.6 kb fragment (including bases 31684-44109 of SEQ ID NO: 27) was purified by QIAEXII (QIAGEN).

[0249] On the other hand, the plasmid vector pSB200 was completely digested with NotI and then DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then dephosphorylated by CIAP (TAKARA). The reaction solution was electrophoresed on an agarose gel, and then a vector fragment was purified from the gel using QIAEXII (QIAGEN).

[0250] The two fragments prepared above, i.e. the 12.6 kb fragment from XSG22 and the vector fragment were subjected to a ligation reaction using DNA Ligation Kit Ver. 1 (TAKARA). Subsequently, transformed plants were prepared and studied according to the method described in Example 4.

(Results and Discussion)

[0251] All of the 48 transformed individuals were sterile. This indicates that the 12.6 kb insert fragment does not contain at least the full-length Rf-1 gene.

Example 10: (1) Complementation assay for a 15.7 kb fragment from XSG16

(Materials and Methods)

[0252] The λ phage clone XSG16 (Figs. 1 and 5) was partially digested with NotI and electrophoresed on an agarose gel. The separated 15.7 kb fragment (including bases 38538-54123 of SEQ ID NO: 27) was purified by QIAEXII (QIAGEN).

[0253] On the other hand, the plasmid vector pSB200 was completely digested with NotI and then DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then dephosphorylated by CIAP (TAKARA). The reaction solution was electrophoresed on an agarose gel, and then a vector fragment was purified from the gel using QIAEXII (QIAGEN).

[0254] The two fragments prepared above, i.e. the 15.7 kb fragment from XSG16 and the vector fragment were subjected to a ligation reaction using DNA Ligation Kit Ver. 1 (TAKARA). Subsequently, transformed plants were prepared and studied according to the method described in Example 4.

(Results and Discussion)

[0255] Of the 47 transformed individuals, at least 37 individuals clearly restored fertility (Fig. 6). This indicates that 15586 bases (bases 38538-54123 of SEQ ID NO: 27) derived from rice (IR24) in the 15.7 kb insert fragment include the full-length Rf-1 gene.

(2) Complementation assay for an internal 11.4 kb fragment in XSG16

(Materials and Methods)

[0256] The λ phage clone XSG16 was completely digested with AlwNI and BsiWI and then DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then blunted by DNA Blunting Kit (TAKARA). The reaction solution was electrophoresed on an agarose gel to separate a 11.4 kb fragment, which was purified by QIAEXII (QIAGEN).

[0257] The plasmid vector pSB11 (Komari et al. Plant Journal, 1996) was completely digested with SmaI and then DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then dephosphorylated by CIAP (TAKARA). The reaction solution was electrophoresed on an agarose gel, and then a vector fragment was purified from the gel using QIAEXII (QIAGEN).

[0258] The two fragments prepared above were subjected to a ligation reaction using DNA Ligation Kit Ver. 1 (TAKARA). After the reaction, DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in pure water (prepared by a Millipore system) and then mixed with *E. coli* DH5a cells, and the mixture was electroporated. After electroporation, the solution was cultured with shaking in LB medium (37°C, 1 hr) and then plated on an LB plate containing spectinomycin and warmed (37°C, 16 hr). Plasmids were isolated from 14 of the resulting colonies, and their restriction enzyme fragment length patterns and boundary base sequences were analyzed to select desired *E. coli* cells.

[0259] The *E. coli* cells selected above were used for triparental mating with the *Agrobacterium tumefaciens* strain LBA4404/pSB4U (Takakura et al., Japanese Patent Application No. 2001-269982 (WO02/019803 A1)) and the helper *E. coli* strain HB101/pRK2013 (Ditta et al., 1980) according to the method of Ditta et al. (1980). Plasmids were isolated from 12 of the colonies formed on an AB plate containing spectinomycin and their restriction enzyme fragment length patterns were analyzed to select desired *Agrobacterium* cells.

[0260] The *Agrobacterium* cells selected above were used to transform MS Koshihikari (having BT cytoplasm and a nucleus gene substantially identical to Koshihikari) according to the method of Hiei et al. (1994). Necessary immature seeds of MS Koshihikari for transformation can be prepared by pollinating MS Koshihikari with Koshihikari.

[0261] Transformed plants were transferred to a greenhouse under long-day conditions after acclimation. 120 individuals grown to a stage suitable for transplantation were transplanted into 1/5000a Wagner pots (4 individuals/pot), and transferred into a greenhouse under short-day conditions about one month after transplantation. About one month after heading, one typical ear was sampled from each plant to evaluate seed fertility (the percentage of fertile paddies to total paddies).

(Results and Discussion)

[0262] Of the 120 transformed individuals, 59 individuals showed seed fertility of 10% or more, among which 19 individuals showed seed fertility of 70% or more. This indicates that the 11.4 kb insert fragment (bases 42357-53743 of SEQ ID NO: 27) contains an essential Rf-1 gene region for expressing a fertility restoring function.

(3) Complementation assay for an internal 6.8 kb fragment in XSG16

(Materials and Methods)

[0263] The λ phage clone XSG16 was completely digested with HpaI and AlwNI and electrophoresed on an agarose gel. The separated 6.8 kb fragment was purified by QIAEXII (QIAGEN).

[0264] The subsequent procedures including the preparation of the plasmid vector pSB11 were performed according to the method in (2) above.

(Results and Discussion)

[0265] Of the 120 transformed individuals, 67 individuals showed seed fertility of 10% or more, among which 26 individuals showed seed fertility of 70% or more. This indicates that the 6.8 kb insert fragment (bases 42132-48883 of SEQ ID NO: 27) contains an essential Rf-1 gene region for expressing a fertility restoring function.

Example 11: Complementation assay for a 16.9 kb fragment from XSG8

(Materials and Methods)

[0266] The λ phage clone XSG8 (Figs. 1 and 5) was completely digested with NotI and electrophoresed on an agarose gel. The separated 16.9 kb fragment (including bases 46558-63364 of SEQ ID NO: 27) was purified by QIAEXII (QIAGEN).

[0267] On the other hand, the plasmid vector pSB200 was completely digested with NotI and then DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then dephosphorylated by CIAP (TAKARA). The reaction solution was electrophoresed on an agarose gel, and then a vector fragment was purified from the gel using QIAEXII (QIAGEN).

[0268] The two fragments prepared above, i.e. the 16.9 kb fragment from XSG8 and the vector fragment were subjected to a ligation reaction using DNA Ligation Kit Ver. 1 (TAKARA). Subsequently, transformed individuals were prepared and studied according to the method described in Example 4.

(Results and Discussion)

[0269] All of the 48 transformed individuals were sterile. This indicates that the 16.9 kb insert fragment does not contain at least the full-length Rf-1 gene.

Example 12: Complementation assay for a 20.0 kb fragment from XSH18

(Materials and Methods)

[0270] The λ phage clone XSH18 (Figs. 1 and 5) was completely digested with NotI and electrophoresed on an agarose gel. The separated 20.0 kb fragment (including bases 56409-76363 of SEQ ID NO: 27) was purified by QIAEXII (QIAGEN).

[0271] On the other hand, the plasmid vector pSB200 was completely digested with NotI and then DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then dephosphorylated by CIAP (TAKARA). The reaction solution was electrophoresed on an agarose gel, and then a vector fragment was purified from the gel using QIAEXII (QIAGEN).

[0272] The two fragments prepared above, i.e. the 20.0 kb fragment from XSH18 and the vector fragment were subjected to a ligation reaction using DNA Ligation Kit Ver. 1 (TAKARA). Subsequently, transformed individuals were prepared and studied according to the method described in Example 4.

(Results and Discussion)

[0273] All of the 44 transformed individuals were sterile. This indicates that the 20.0 kb insert fragment does not contain at least the full-length Rf-1 gene.

Example 13: Complementation assay for a 19.7 kb fragment from an overlapping region of XSG8 and XSH18

(Materials and Methods)

[0274] A plasmid (XSG8SB200F) isolated from desired E. coli cells obtained by ligation in Example 11 was completely digested with Sall and StuI and electrophoresed on an agarose gel. The separated 12.8 kb fragment (including bases 50430-63197 of SEQ ID NO: 27) was purified by QIAEXII (QIAGEN).

[0275] On the other hand, a plasmid (XSH18SB200R) isolated from desired E. coli cells obtained by ligation in Example 12 was completely digested with Sall, StuI and XhoI and electrophoresed on an agarose gel to separate a 6.9 kb fragment (including bases 63194-70116 of SEQ ID NO: 27), which was purified by QIAEXII (QIAGEN).

[0276] Further, the plasmid vector pSB200 was completely digested with EcoRV and then DNA was recovered by ethanol precipitation. The recovered DNA was dissolved in TE solution and then dephosphorylated by CIAP (TAKARA). The reaction solution was electrophoresed on an agarose gel, and then a vector fragment was purified from the gel using QIAEXII (QIAGEN).

[0277] The three fragments prepared above, i.e. the 12.8 kb fragment from XSG8, the 6.9 kb fragment from XSH18 and the vector fragment were subjected to a ligation reaction using DNA Ligation Kit Ver. 1 (TAKARA). The ligation product contains a 19.7 kb fragment from an overlapping region of XSG8 and XSH18 (including 50430-70116 of SEQ ID NO: 27) (XSX1 in Fig. 5). Subsequently, transformed individuals were prepared and studied according to the method

described in Example 4.

(Results and Discussion)

5 [0278] All of the 40 transformed individuals were sterile. This indicates that the 19.7 kb insert fragment does not contain at least the full-length Rf-1 gene.

10

15

20

25

30

35

40

45

50

55

SEQUENCE LISTING

<110> JAPAN TOBACCO INC.

Syngenta Limited

<120> A method for providing and inhibiting the rice fertility, and discerning the presence of the rice restorer gene by using the rice restorer gene to the rice BT type cytoplasmic male sterility.

<130> YCT739

<150> JP 2001-285247

<151> 2001-09-19

<150> JP 2001-309135

<151> 2001-10-04

<150> JP 2002-185709

<151> 2002-06-26

<160> 68

<210> 1

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of R1877 EcoRI marker sequence.

<400> 1

catctcctgct tccatggaaa cgtc 24

<210> 2

<211> 33

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of R1877 EcoRI marker sequence.

<400> 2

ctctttctgt atacttgagc ttgacatct gac 33

<210> 3

<211> 20

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of G4003 HindIII marker sequence.

<400> 3

gacgcacgag tacctgaacg 20

<210> 4

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of G4003 HindIII marker sequence.

<400> 4

aatagttgga ttgtcctcaa aggg 24

<210> 5

<211> 27

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of C1361 MwoI marker sequence.

<400> 5

aaagcaaccg acttcagtag catcacc 27

<210> 6

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of C1361 MwoI marker sequence.

<400> 6

ctggacttca tttccctgca gaggc 24

<210> 7

<211> 27

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of G2155 MwoI marker sequence.

<400> 7

gaccaccaat taactgatta agctggc 27

5 <210> 8
 <211> 27
 <212> DNA
 <213> artificial sequence
 <220>
 10 <223> Oligonucleotide primer for amplification of G2155 MwoI marker
 sequence.
 15 <400> 8
 ttctggctc caataatcag ctgtagc 27
 20 <210> 9
 <211> 27
 <212> DNA
 25 <213> artificial sequence
 <220>
 <223> Oligonucleotide primer for amplification of G291 MspI marker
 30 sequence.<400> 9
 ctgcigcagc aagctgcacc gaaccgg 27
 35 <210> 10
 <211> 24
 <212> DNA
 40 <213> artificial sequence
 <220>
 <223> Oligonucleotide primer for amplification of G291 MspI marker
 45 sequence.<400> 10
 acattttttc ttccgaaact tccg 24
 50 <210> 11

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of R2303 BslI marker sequence.

<400> 11

atggaaagat acactagaat gagc 24

<210> 12

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of R2303 BslI marker sequence.

<400> 12

atcttatata gtggcaggaa agcc 24

<210> 13

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of S10019 BstUI marker sequence.

<400> 13

aacaatccta tcctgcacag acig 24

<210> 14

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of S10019 BstUI marker sequence.

<400> 14

gtcacataga agcagatggg ttcc 24

<210> 15

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of S10602 KpnI marker sequence.

<400> 15

agctgttgag agttctatgc cacc 24

<210> 16

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of S10602 KpnI marker sequence.

<400> 16

tagccatgca acaagatgtc atac 24

<210> 17

<211> 26

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of S12564 Tsp509I marker sequence.

<400> 17

ctagttagac cgaataactg aggttc 26

<210> 18

<211> 27

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification of S12564 Tsp509I marker sequence.

<400> 18

tttgtgggtt tgtggcattg agaaaat 27

<210> 19

<211> 2240

<212> DNA

<213> Oryza sativa L.

<223> PCR marker G4003 HindIII

<400> 19

gcggccgcctc cgggaagtcg agcgagtaga cgccccctgac gccgtacgcg tcggcgagcc 60

gcagcggcgt ctcctggcgtt gigaaggaca gcccgcttcag cgtcgcgcgg cgccgcccgt 120

1gacggtcac cggcgccgtg ctccgcagca ggtacgccig cgtcacgltg atcgacgagt 180
 5 acclgaacga lccctglggg ttcggccctcg ccgcctccggc actcagggtc caccigccca 240
 algcacaaaa ccaaacacca aaagccta at gcaataata catcattcca cgtattttaaa 300
 aaaaataatt ataggtaaaa tttttataat gtatttttagc gacgtaaatg tcaatgctga 360
 10 gaaataaacg ataactttt aaatgaagt ctacaaattt aattttggca tgggttgatg 420
 ttggataaag aaaacgatgg aggcctagtaa ttttcttct tttttaagta tctagattgt 480
 cataatattga atttttcagt ttctcatccc ttgaggaca atccaactat tattttcctt 540
 15 ttcttatgta aaagggtgaa caacatattc aaacataaaa aaataaaaatt aaatgaaata 600
 aatttacaat tcataaaaatt tacagaattt atgttaagaa aatattcaaa cttagataat 660
 aataaagcaa caaaatcgta ctacaaagaa gtataattgt acattgtata ctactactcc 720
 20 tacaatttta gacttagaat ttttaatttc ctgaaacta gtaatgccat tttttcttt 780
 ctagttagaac cagacagtaa gttaactcg aaacttataa gctaattgagc gaagtcgggc 840
 25 aattcacctg taccigacgg agcgagcttg gtcatggag aaggacttgt cgaactggtc 900
 ctggggaggg tgggggagcg ggccggagcg ccgcccccg gagttggagt agcggaggac 960
 ggcgacgccg gcgacggcg gccacacggt gtcgttcacc atgcgcgcgc tggcgacgac 1020
 30 gtagtagtcg gagctcgct tctggctcgt ggtgacgagg aaggagtagg actggccgac 1080
 gtggacgtcc aggttggtgt agttctgtc cgtcgtgtag gagcccccg tctccaccag 1140
 caccaigtg tgccttgga tctgaagt gaggcctcgt gacgtcccca cgttgtgcac 1200
 35 tggatcctg tacgtcttg cgtgtcccc acaccgagc cgcgcacaca cgcgcaaaag 1260
 ataatagact cattgtaagt aggtagtaac ctctccgtt tcataattata aatcgttga 1320
 ttataatttt gttagtata ctcttttaag tttttttct ataaacttaa ttaaacttaa 1380
 40 agaattttta taaaaaaat caaacgact ataataaaa atggaaggag tagttgcatc 1440
 aatttgtgga tgaagcaaac aagattatat cttttcatg agggtgaaag tattcagtga 1500
 acaattcgtc agtttcaagt ttcatgaaat cggacagggt ctctgaaagt ctgtattttt 1560
 45 ggtactgttg gattgactac tctggttct gtgtgacat cttttgtatc ctagtctcgg 1620
 taaaaaaaat ttggcattt ttactctat cgttgatctg tttaactgaa accattgcat 1680
 galatactac tagcagacaa aactggtgaa aattcacgag aatgaacttt ttgtcagtt 1740
 50 agcattagcg gacagcttca gtaagcagag caggctgcct taaggcttaa agcactatct 1800
 tccacaacac ttgttctac aatcaaatc caaatttact atcacaaaa gcgaaggaa 1860

taactaaacc ttactcctac tagtactact gctatgacta tgaaacaaga ttccaatcca 1920
 5 aagaaaacac agtgctcgat cagcatgata aaagcaacga aacctgctca tccagctgcc 1980
 aaaatgccac cccactgact ctacgtacgt actacgtatt gacgctgtaa aaaactagcc 2040
 gtagtacaga gaagaggacc caaagtittc tcaaaaattt tatttlaccc ggaicccat 2100
 10 tgatggcttc gtaactgatg cgggccggga caaggctgtc gtigtacctg tacgggccc 2160
 tgccgtaat cagcacgccg tccggcatcc cgaggctctt gccactgtcc agcatcttcc 2220
 tcagatcctg caacgaattc 2240

15 <210> 20

<211> 2601

20 <212> DNA

<213> Oryza sativa L.

25 <223> PCR marker C1361 MwoI

<400> 20

tcttgctgag atccaagttg cggtaacttt gcccttttct tttttcttc tcttctgaat 60
 30 tttttcatgg tttttgggag agattttcgt aacttgatta cagtcttagg aaaaggccac 120
 ctgtttcaaa cagggccttc ttgaaaggga tcaatttgct aggagtaaat gattctaaaa 180
 gcgatttcga aataaaacac agttctcgat ctcataccig aaaacaaaag gcccatactg 240
 35 tglaaactgt gattatgctt ctgtttaaag ggatatttgt acaaaattga cgccaaccac 300
 ctataaacag atttgtagct ttatcttag taaaataaaa tggacattc tactcagigt 360
 tcagtgatcc gatgctgctt ctctgctgta caacttctaa cagccgtttt cggtagtaca 420
 40 aactagcgaa acacaaaaa cgcagcatit gatttctgga atacgctgaa attgttagaa 480
 tcaaccacga aacaaaaatc attgttcaga aacgttgcaa cgagataaaa cacaagaact 540
 tgttttaaca aagcatacgg acagtacata tacggttaca acaccagtc ttatacagt 600
 45 tctgctggag ttccatctac tggctgtcat tgtatctcag gacagacagg ttaacatagg 660
 tacaacacaa ttacaggcta aaccgaagcg aactacactg tcagcatctc taacagtatc 720
 gtaagcaag ctattttaca gctgctctag taaatttaca acgtccctgg cagaatccct 780
 50 ctgctttctg gcagcgacga ggcacggtcc atggccttag caggacatct caccgctcag 840
 ctgcatagaa agcaaccgac ttcagtgga tcaactcctg ctcttgcaaa aaagttaggtt 900

cgatcaatca cgcgtttaa ccaaaacaaa atggglaata attatgctag cctatgaagc 960
 5 tacctcagag ttctctatctt gctctgcagg gaaatgaagt ccagtggaac agttctcaag 1020
 cacctcaggg ctcttcatcc atgctttgtg tgcctcaatg gctttcagct tatagcgaaa 1080
 catctgcgat acggaactaa aattaaggat gtgcacaatt acttaacaca acaataaatt 1140
 10 gaagcaggtc cagttaaaga aaagtagcag cgaagaatag cactcgaag tctgaacctc 1200
 agataaagaa atgggtgggt ttccagttc atctccctca acatggattc cagtaccttg 1260
 gcattctggg caaaggatgg atgttatctt cttagggtgca ttttttgcct ttcttctctg 1320
 15 attgcttttt cctttgcttg caattttgtc tgcctagcatc tcataattggc ataaaaatgt 1380
 ccagtgacac aggcagaag tctgaacaaa atgaaatgcc tgcataaata gccgtacaaa 1440
 gtcattggag gtgcagcag aatactacaa atttttaaag aagaacctat acactgtcta 1500
 20 tgttttgcct gaaatgaatt caaccacttt gcattatagc gtttggaatc cctggtttgt 1560
 gagaactgta attccattac aacagtgaag aagttacct aactaatgaa tggaaattag 1620
 tcaaatgcct aattttttag gtttgcctta atttatctat ctgtgagaaa tgcataagcat 1680
 25 gtcattgcgt gctatcttca agaaatacta agaaactgca aaggcaaaga atgtttgaaa 1740
 taacttacct cgtttgagtt tctactgtg caggctagat ttctgtctt gcagttgagc 1800
 30 aaggtagcta catcttttc aagaagcatt ggtcgccac aaataacaca agctttctca 1860
 gcagcaaggc gcttctgctt acgcaactcc ctctcatag atttggtgga taagaggcca 1920
 acttgaagat tgttgaagt acctgtcggg gaacctgtta tgatagcttg gctattgtca 1980
 35 tgggcggagc tgccttgctc attcgactcc tctgaagatg ctctttgatc tgaataatgac 2040
 ttctttcttc tcttccacg ggtccagca tcatcaatca cgaagaaaga tccagcagag 2100
 ataggaaggt cctgatcatc agaagaccac ttctgtccca actcaattgt ataagagaag 2160
 40 ttgacaatgg caaagtcaga ttgtcatag gtgtcacact catccaagcc atgggagcca 2220
 tctgttctta cccaagcaca ccagatcttg ctaatcttt tacttcttt gctagcttcc 2280
 cataacctgt atgcaatatt tccatatccc aaaagaatgca caggcaaata cgaaacaaca 2340
 45 tcttttagca atacactagg aataacgaga ggaccgtcag ttccactttg gtttgacagc 2400
 acatgatctt cagatacaga agcagttcta ccattacct gcgcatttgc accacggcgt 2460
 50 gtgctttttg cgccattgcg agagctagaa tcatcttca acctcgaagt cacttcagt 2520
 tcttctgtg gaaccagagc cagctctctg ggttctgcg agctcgagtc cagcaagagc 2580
 ggttcttct cgcgcgagtt g 2601

<210> 21

<211> 1333

<212> DNA

<213> *Oryza sativa* L.

<223> PCR marker G2155 MwoI

<400> 21

ccctctgctt gatccagtgt acatccatgg gttaggacag attagttact cagttaatta 60
 agtgtgagac tggaaaaaaaa taictgacgg cagttttata agttgagtga tgaactagtt 120
 gaaagttcag ttaactgtca acggctgtag atttgggaig gcagactgtt ctgagtcaaa 180
 atgaagcitt tactgtgcgt ggltaccagg tgcagtaaaa taatttcaga tctaattgca 240
 gtaaaaaaat glagtactat atglttaagac gagattgggtc ggtcaaaaac tatctggccc 300
 ttltacatctc ccaaatgtta cctcagttgc aggtggtaaa aaaaaatcac tctgttcacg 360
 tgaatgtcggc agatcatgga ccaigtctca aatgctgaaa ctctgaacaa tcaacaaaaa 420
 aatccaacca gatgagctgt gcaactgata attgatcatc acactatttg caactcatct 480
 ttcatgtaga tggaaactta atcccgaaga aataatgaca gcaaaatgct gcgatccatga 540
 agaaaggatg gcggcaaaat ggcagcgata aaaaaaaaaat ggttggttac tgaagaatta 600
 tttgtgcagc agttgagaca glagcaagat aagagctagc taagctagct aggttagagtt 660
 ggatggaaga gtagtagtat gagatagagc atggagcgcg acaactcaag tggatgclaa 720
 agtaaaaggc attctcttct ctigtittgga atcagaaaag aaaagaaaag acttgagctg 780
 ctltggctgga atgtttgggtt ggalcatgcg cgctctccct agcttagctc gccaaagaaat 840
 cctcgcttca tctctctcaa taattcaaag ccacgagctc tctgctcata tccagtgcga 900
 cgattcccgt taatgcaaat gcattatctc cagttcgaaa tgttacaatt ctltgcgtttg 960
 cagcaagcca gcaagtgggtg tgaattgttt aatccctcgt gcatttcaac gaaattctct 1020
 cacaaattcg cattgacttc ttcttttagca caattagtaa gcagtgacaa ataaagaatt 1080
 ttigaacagg atgtctttcc aaggaagggt agatttttta tgttgatagc aaggatcgcc 1140
 ttctcttagc atgaagagaa tgtgatcaac ttacacctt gcttacgatt atggccttaa 1200
 tttttgatac cctaaacagg agcacatcac atgcatgtcg acctgagacc accaattaac 1260
 tgattaaagt ggcatctcag atgcatccgt cagttacatg atcaggtagt cgatggatca 1320

actgtaggtt tca

1333

5

<210> 22

<211> 863

10

<212> DNA

<213> Oryza sativa L.

<223> PCR marker G291 MspI

15

<400> 22

cgaacaggat caaaagtaga cgacgagggc atttagaagg agaggaattg tatttggtcc 60

20

cggtatttaa tttttaaatl igtggtcgga agtttcggaa gaaaaaatgt gctcatgagt 120

gattattggc tctgaacacc aaccctcctt ttcggtgatt ccttcigagg tgttgggtgt 180

tgggacacga tgcigccgcc gacacgacac cgggttccac aalacactaa tctactcgcg 240

25

acaccitcat tgaacigcat ataattatit agaaagtcga ttaacacatc ttataaaacc 300

ttgttgaatc atataatcat tctataaagt ctatttgaac atcttatgaa aaaataagat 360

ctgacctagt cgttacacac tcttacattt tccattagcc taactaatc cgigcaggaa 420

30

acgccccaaa ataatagtac caatagtcga ctaatcccgt gccagaggcc gccaatgatt 480

agtgattaac ccaaaaaaca taatcatcat cacacgccgc taatgaccag ctctcgctta 540

gcicatccca caggcgggccc ccacacgcca ctcttgccat gtggggcccac ctttcacacc 600

35

ccccaccaac cagaaaaaaa actcccccaa aaaaaaaact tttaatgctt atctcgcggc 660

agtataaaaag gcgacccccc cacccacaca caatcacagt cagcgaccca acccaacccg 720

agccgaggag tcgagtcgtg tgaaaattac gaaattgccc ttcgactcca ccaccaccac 780

40

ccaccggcga ggcgaggaga ggagaaaaat tgggaggaaa aaaaaaggga aaaagaaaaa 840

gggtggagga gatttttgcg aag 863

45

<210> 23

<211> 1510

50

<212> DNA

<213> Oryza sativa L.

<223> PCR marker R2303 BslI

55

<400> 23

5	tgccatgaag acctatggaa agaalatctt ctctcactc tgtgaatggt gagtctactc	60
	tctgtaacat ttagggttag gtcgaaggaa catgaagcat tgcgtattca ctccactgtg	120
	tttttttttt ctgtataggg ggaaagaaaa tctgtctaca tgggcaggcc gcatgggtaa	180
10	cagctggaga acaactggcg acatcgccga caactggggc aggttctact catcctctct	240
	ttaacccctgt ttacatagtt cttagatttt tcagttactga tctaatgtc cctgttattt	300
	cagttatgaca tctcgtgcag acgaaaatga ccaatgggct gcctatgctg gacctgggtgg	360
15	atgggaatggt aagaacttga gatgtatctg ttccctagggt gcttaacat ttgagagctt	420
	caaaatgatc aacatatgtt tctcgtgtgc aatatcagat cctgacatgc ttgaagtggg	480
20	aaatgggtggg atgtctgaag ctgagttaccg gtcacacttc agtatctggg cactagcaaa	540
	ggtaccatag catgttctat gtactaataa ttttgcctga atgttgaact tctttgcatt	600
	tcttcactgc aagttttgt tgaattgttc aggctccctt ttgatcgga tgcgatgtgc	660
25	gttcaatgag ccagcagacg aagaacatac tcagcaactc ggaggtgatc gctgtcaacc	720
	aaggcaagcc ttctcagitt cacatgctta gatttagcca tactctctgg atatttcacc	780
	alactcataa tgtaaccttc tgaacagata gtctagggtt ccaaggaaag aaagtacaat	840
30	ctgacaacgg attggaggta tcccttcaat ggcttccaaa ttgacagttt ctcatgttcc	900
	cataagcctt ggcatgatca tgactaactc tgaagctgac aatactttgt glaaaattgt	960
	cggtaggttt gggccgggccc actcagcaac aacaggaagg ctgtgggtgt ctggaacagg	1020
35	cagtcatacc aggcaaccat cactgcacat tggctgaaca tggggctcgc tggatcggtc	1080
	gcggtcactg ctctgtatct atgggcggta aagcctttgc ttctttcaga gctcaaagta	1140
40	gaacatcttc tcttcagaat tcagagtcca taacaaattt ctgtcaattg tgcagcactc	1200
	ttcgttccgc gctcagggac agatatcagc atcggttggcg cctcatgact gcaagatgta	1260
	tgtcttgaca ccaaactagt cagcaaagaa aagcagcaca ggttagttacg tgtccggcga	1320
45	alacagctaa attgatcagg attcaggaag aaggtttgca atttgcaagg attggtagag	1380
	ctggaaatgg gatgccattt ggttatgtat gttagaaataa gctgttaagcc tgttaagcgt	1440
	tatgtaatca gccgtcaaat gctggcgagt gtatttctga agtttgcaac gaaagttgca	1500
50	gcaataaaaa	1510

<210> 24

55

<211> 1016

<212> DNA

<213> *Oryza sativa* L.

<223> PCR marker BstUI

<400> 24

```

tggggattct tttctttaag caatttaaca ttattgtcct aacaatatac acaatattgg   60
ttttcttttc agtatcaaat aattctttta cttttgaaaa cacatttgca atgtgttgga   120
aacacaatta tatcttgcac ttcttttgg aaatttaatc atttgaaaac tgattcgcgt   180
ttcatggctg taatcttctc ttgcgaacat cgcctttctt ttgatggttc tctgttgaga   240
agaagagcaa ccaagtaaat tticgaaatg tttttttgtt ctttctattc accattgcag   300
gttgtcaaaag ccatcgagaa ggccataccg attccgagag cgcaacccat tgccttggat   360
ggcccagcaa gggaagagct gaaggccatg gaggcgcaga aggtcgagat cgaccgcacc   420
gcggcgctcc aggtgcgccg tgagctttgg ctggggctgg catacctcgt cgtccagact   480
gccggcttca tgaggctcac attctgggag ctctcatggg atgtcatgga acccatctgc   540
ttctatgtga cctccatgta ctcatggcc ggctacacct tcttctccg gaccaagaag   600
gagccctcct tcgagggcct cticgagagc cggttcgagg cgaagcagaa gcggttgatg   660
cacgcccggg atttcgatct ccgccggtat gacgagctcc ggcgagccctg tggcctgccg   720
gtggttcgga ctccgacgag cccctgcaga ccgtcgtcgt cgtcgtcgtc gtcttcgacg   780
caggagagcc attgccattc ttactgccat tgccaatgat ctttgtgtctg ttctgttctg   840
ttgtcagaat tttttcatgc ccagtttatg ggggttaagc tagcttctcc attgtaccgt   900
tcgatgtgc ggatgatgcg atgcaaagca tagtttgttg aagagatgac aaggcagatt   960
ttagcttgaa aacctggagg tgagaaaaaa aaatcctgat gtgttttgtg gtgtga   1016

```

<210> 25

<211> 676

<212> DNA

<213> *Oryza sativa* L.

<223> PCR marker S10602 KpnI

<400> 25

5 accaccitca tatgaagaaa ttaacgggtgt tticatgagg aatccaacag tgcctgaatt 60
 ggiggaaact giggaaatct tcttggccta gglaaccaat catcacttca ccacaatgca 120
 caagtttcta gcttactact acagttactt taataagttt tgtctgttga gattttattg 180
 ctgatttcta tgcattggca tctttttgac aggccatcca gtcttatcgt gctgagagtg 240
 10 aaactgagct caacctggca gctgggtact atatagtgtt ccggaaggta cggccctatc 300
 tcccattgg acatgttctt aaccataaac atattcttgc tggacttttg tgggcaaagt 360
 tggctacact aaactgtgtt tcatlaacct gctcaatcag ggttcaaaca atggatgggc 420
 15 agaaggtaga tgcagaggga aagctggctg gtcccttacc gactacatcg agaaaaggga 480
 ccgtgtgtctt gcaagtaaaag tgcctcagggt ctcttaggcg ttcaatgagc catacataca 540
 taaccttgggt gtgttacact gtattatgat cgttcgtgat ctcaaagac cctctgatca 600
 20 gagaaatcac aaatattctt ttgttctatt atgttcatla tcaclacccc ttttgtcaaa 660
 accagtgcag cctttt 676

25 <210> 26
 <211> 1059
 30 <212> DNA
 <213> *Oryza sativa* L.
 <223> PCR marker Tsp509I
 35 <400> 26

gcgagatcat gaacttgatt ttctgggtgc catattgggc ttgtttgta accttgtaga 60
 gaaggatagc cttaalaggt aagtccttca catgttctt tccatttgc caattcataat 120
 40 cagtgttact gtcttggcag ttcttgggg tcaggactca gaaacatcca attaatgttc 180
 atgttctctt aacgactcag aaatacttta taacctctcc acagggtacg gctttcatct 240
 gcccgtgttc ctgttgatct atctcagaat ccacagagtg aagagacaca gagagatgtc 300
 45 atagcactcc tctgttctgt attcttagca agtcaagggt ctagtgaagc ttcttgaact 360
 alataccggg taattcaaaa ttcttcaagt tcttttcta tgtagattat atctttgtaa 420
 aactcggcat ttattacctg ctctttgttt caaaaagcag tattttattt tgctccctag 480
 50 cataggctcag cagaacagtt gatcttattc agaaaacaat attttgcatg taacatactg 540
 ttatctatga gatgaaaatt aatgcatgtg taataatgtc aatgataaat atttgctaic 600

5 tgaatccagt ctaccaactc tagttagacc gaattactga ggcttctatit caaagaataa 660
 ttttagtgcac catttgttca actactatga agtaaaatgg tattcccttc tatlgacatc 720
 gggttagaag tgaaaggcca tcttaatgcg atgttctcaa tgcacaaaac ccacaaattt 780
 cattaacaca tacagattat tattaacata gclataaatt ggatttccag aagcttgagt 840
 10 tgaatttatt ttgttacaat tgaaagcact gggaacatta gcattttttt ttagttcttg 900
 gttattgcaa ttataatgt tatacagaac tgggtaccic acaatgcatt cattatgaca 960
 ttctatgaac catttgattg acgtttgctt gtaacaaca ggatgatgag gagtcttga 1020
 15 tgcaaggagc acgggaagct gaaatgatga tcttagagg 1059

<210> 27

<211> 76363

<212> DNA

<213> Orza sativa IR24

<400> 27

gatcaactaa caacctcttt gcagcaaaaa agcatacaca caagtgtttg tcttggccctg 60
 30 gggctctgca gatggaciga tactctgacc tgcagtgggc ttgggagcta acaatggttt 120
 catctctttt ttttttatgt tttccctgt tgttttgcct catgtttgti gtaattttt 180
 ctctctatct agcgaigta ttttcttag catgaiggga gtagccctcc ttttttttc 240
 35 tctaatlaag tgtaaagtag caacagcata gggatgaatg ttcagttag tgtgtggtgt 300
 ttcagttatt cagagacgtc calacagttt giacctgtg accacacgtc ttaatctgat 360
 gaagcttaga ataaatcaca tgttagcaat gcaatatcat ctgcgtcttc tctcacttg 420
 40 gttggccatca aattctgtgt agaagtgtat ggttggigtg ctgttgcaa tgcctgattc 480
 cgtctctgtt tgtggaagtt aagaagctcc tagttgaaat accgatitit catgatctcg 540
 gagattgatg caactctgat tgcagcattt ctttttata gaatgtacac tccatgclat 600
 45 catgatgttt attgtttagt actacaagat ttggtlaacc attattttta tatcataata 660
 attttataaa atcttggagt acaagttca taatacatga tagcataact ttttgaggct 720
 50 agtctatgta tattgtctcc ttgttttita aactaagcac tcaataaatt attgatggct 780
 glaattttct gaaggttica ccggtttcgg cccgtgcttt ataaatagct tccgcacaaa 840
 agacaaaacg gtccctccaa cacataaalg gttagttita cgttttcatt atcttttgta 900

aaatcaagtc caccacgtag acactcataa caaaagttg aatatccetca gaaattttga 960
 5 cttgagtcia tcttaccitl galatcggac atccaaccct ccttccctcc ctgaacttta 1020
 tattattcat attacaccig aactttalat lattcatatt acaccctgaa gtggttttca 1080
 ttttaattgca tacatgciga aatagttlga caacgtgaga tgcactaaaa atctacacgt 1140
 10 tegtcttaag ttgcaattca ttttatccct tttctttttc tctcttacct aggaatatca 1200
 atagtactaa ttcacattac aatatagtat aaattggtaa tggattatlg gcaatatact 1260
 atattaaata ttcaaaacta gtcatttaag ctgccaataa aglaaaccac tatcgaaaac 1320
 15 cacaalataa atggcattac aaaacttagg gggttgaata tccaatttta aagttcatga 1380
 tgctagagga atttctatca aaagtittaig ggtacataig gactttttcc tttttaaaag 1440
 aagctattct tgtcgtaaac gttaaatatt ttttgtactt ttttttttat gattgaaaaa 1500
 20 aaaacttagt ttcaaaaatg attggtctgt atacaagcat caattagact taataaatlc 1560
 atctaacagt ttccggcag aaactglaat ttgtttlgt lattagacta cglttattat 1620
 25 ttcaaatatg tgtacgtata tctgatgiga caaccaaac caaaaatttt ccttaactcc 1680
 atgaggccit acagalatat ttgatgggtg taaagtttt taagttcttt gggtgcaaag 1740
 tttttaagat atacggacac acattigaag tattaaatat agacaaataa caaacatat 1800
 30 tacatatctt gccigtlaac aacgagacaa atttattag cctaattaat ctgtcattag 1860
 caaacgttta ctgcagcatc acattgcaa atcatagcgt aattaggctc aaaaatattc 1920
 gctcgtlaat ttacalgcaa actgtgtaat tggtttttt ttctgcaaca tttaatactc 1980
 35 catgcatgtc caaatatttg atcgatctt ttggccaaa tttgttggga atctaaacaa 2040
 ggalcaaatt tctgaattt ttccagacgt cagggcttgt tcatccatcg ttcgatcgc 2100
 gattcgccac cgacgccttg gtttccaacg aattttatca tccgcttaaa tacatccaaa 2160
 40 gctctccatc gccatcggcg gccaacggcg accgctccgc tctacceaat ccacccatcc 2220
 actcgccgcc gccccctgat ccaaagcctc cgccgcgccg ccgtcgagag gaggaggagg 2280
 aggaggagga ggaggcgtga gcccctalgg ggacctctt cgggccgctt ccgcttgccc 2340
 45 acgccgccgg cgccggcgac gccacgccgt cgaccgcga cggtagccac gcgctctcg 2400
 agaggcccc ccccccgcc gctcgtgat ctctctctc atctgtttg ggtttgggtt 2460
 50 tglgatttgg gtgtttttt tttttccgca gcggtggtag tgagcggtgg ccgcggccgt 2520
 ggcgtggagt gccagccga tgggtgctc cgccgccgg gtccgcaggt tgcggtggcg 2580
 acggcgagct ggaggaggcg gagggagacc gtggtgagat cggatttcgc cgtgggtgtt 2640

gccgcclacca tgggggattc gccgcaggcg ctctcagggt tgcagcctcc tccactctct 2700
 5 tctcgcaaaa tgtgttgcta tgttctctc gcctgggctgg cctcatagcc attaatgtag 2760
 ttgtciggaa catlacattc ggaacgttgt tggcaattgc ttgacaaaat gtggaattgt 2820
 ggaggggaga aaaatcggtt gaacctgcag tgacaaaatt gccatctata attttaaaac 2880
 10 tgaagggtgtg gaaatcaaac alaatacttg ccagcacatc attcttgtta accaccttga 2940
 catatttgtg gcttataaca gttagctcca caccaacttg gaagggtgtc atggaatgtc 3000
 agtataaatt gaggataact ggcagttgtt aagactttct acagaacttg tagcagctaa 3060
 15 aactagctat tgtgcatlla tgtttcatgg aatttgagcg gcaatggata tttcttacta 3120
 agacglataa tgcaaaaaaa aaaaaaaaaac tatgtctatg cagtttacat gtaatgtgcg 3180
 gatgcaaaata aaatcatgtt catggacaaa ctaatgggat tcalaccaa itccagaatt 3240
 20 gcaittctta tgtggttact ttgtttgtt gatttggtta ccagacatcg atgtggttct 3300
 aagggtcaga ggggtttgtc tctacgcggt gactgcagtt gcagcaatct tttgtttgt 3360
 cgccatggtt gtggttcac cacttgtgtc cctatttgac cgalaccgga ggagagctca 3420
 25 gcactacatt gcaaagattt gggcaactct gacaatttcc atgttctaca agcttgacgt 3480
 cgagggaatg gagaacctgc caccgaatag tagccctgtc gtctatgttg cgaaccatca 3540
 30 gagtttcttg gatactata ccttctaac tctaggaagg tgtttcaagt ttataagcaa 3600
 gacaagtata ttatgttcc caattatttg atgggcaatg tatctcttag gagtaattcc 3660
 ttgtcggcgt atggacagca ggagccagct ggtaaggctg tagtctcatc cctgccttct 3720
 35 taagtagaca tatataatt tacagtattt ggtaaataaa caagatttta tgaatcata 3780
 atgatttttg ggaaaacaca aaactctctt tgttggctgc ctggaacata gttctgttca 3840
 cacagttata gcaccttctt taaaatgaag aactttgttg catcacata aggccaaacc 3900
 40 acataatgaa tttgtttat ttctatctt gaagttagc atcgtttttg tttaatgcat 3960
 gatcgccctc ctatatatt ttagtatgtc aacattgtat tccatgttga gcataacaaa 4020
 tggtttgtta aaattcagga ctgtcttaaa cgggtgtgtg aatttggtgaa aaaaggagca 4080
 45 tctgtatttt tctttccaga ggggactaga agcaaagat gaaagctagg tgcatttaag 4140
 gttcagtaac caacttagg ttacattaca tctaalgaga tttttatatt cagtataaa 4200
 50 tgttaacctt ctcatgggtt actgacgttg ttataaatgt cccagagag gtgcattcag 4260
 tgtggctaca aagaccgtg ctctgtgat acctattact ctctcggga cagggaact 4320
 gatgccttct ggaatggaag gcatccttaa ttcaggttca gtaaagctca ttattacca 4380

tccaaltgaa gggaatgatg ctgagaaatt atgttctgaa gcaaggaagg tgaatagctga 4440
 cactcttatt ctaaacgggt atggagtgc ctaaagaaag atgggtgttt tttttattat 4500
 atggaacctt ticaaaggca cagacaggct ticaaggcta agcttggtac aggtactgat 4560
 actagttact aattactttc gtaatcagta taaataagct tgtgtagtgt aatggcatig 4620
 tacaattctg cacttggtaa atttacagaa gaggaagta atatittaga ggattgagtt 4680
 taltcaccca gtcatatagi tgaagaggca agtaacctgt aagagaggac tgaacattaa 4740
 caccctctgt tcatataaaa atgaccaaag agcatcaaac atgtattcga ggctgttact 4800
 tttagatatgg cccattaaatt tgtttagtgt tcatgtaca tcttagttgg tgtaaatgcc 4860
 agttaccatt tcatgatct aaaacaatca actcttttag tataatttca aaaacgaaat 4920
 tcagtacaca tgaatgaatc ttaataattc tctctagctc gttacaaaag caacaaaggc 4980
 accgtgtcag ctgggtcaca ttagctagtt tgtacttagc attatccact agcaccttat 5040
 tttcaigcat atcatgctaa ttgtcttgc cactgtgagt gggaattttt tcatgtttt 5100
 ataatttata tatgttttag acttctagtc cacaatttat glacttcaig ttcctgagcc 5160
 tctaglatgg ctgatagcag actagggtgt gagtgctgtc cttttttgca gactgaagag 5220
 agaagaaata caagactgtc cattgttagt cagatttgta aaaalagact ctgatgtagt 5280
 ttacttttgc cctatttita tttttaacaa tacaataata taacagatcc taagaactta 5340
 tcttaattta ggagaagttg ctctgttcat taaattaaat tgtgaagtaa aaatgtgtgc 5400
 tggagtctgt caatgcaatc ctgtgttctt gtttgaagat atgggtgagg gcaggccagg 5460
 attgaacact gaatggtaag actgcctctg ccttcagacg ttattgctaa attttttagct 5520
 acttgcagtt agtgctgcca cgccgattaa gcagtagaac aaagtgttt tgtcgtgcac 5580
 aaatgagtta tatttcattg gaaatcgaag cgaaaacgaa tcaaaagtta gaagaaaagg 5640
 ggaaacttgg taattactcc ataaagagag tgcattttat tggtaagatg gtaaccgaa 5700
 gctgtgagct ccgggctgta tgaattctgg caaatttgat atgagatgtc cgattattgg 5760
 ctttaagttag cgataatcaa ttgggggaag caccaaagga attattgtga aggagttaig 5820
 gggtcgtgac gttatctgtc aggttcaaat cctgttggct atgaatattt atctgttagg 5880
 ttc aaatcct agtgactatg aatattaaig ggtaaggtaa gggatttatt gtttaattta 5940
 gtttctttta gatgtgtcca tggacgcca ttcggttaact gtaataatgc ttgtatttgg 6000
 attcacttgt gttacatgca cgcactaaac atgtgcctta ccttttcatc tgtttttgcg 6060
 ttctgggcta gaaactcaaa cgttgaattt tccatgtct gctcaacttg acaattactg 6120

cgtgtcaagc gatcctatag gcatactatg cgcacaagtg attgtatagc gatatgatga 6180
 cagtataacg tggatattg atttttttaa taaaaaaatg atgttcattt ccttgatgaa 6240
 5 ggaacaaaga ctttttttaa aagaagggtt ttactaaaaa caaaaatgac aaaaacaaaa 6300
 tatcagtgcg catggcaagt gtgtcggcga attttttttc tgtactttaa aaaaaatagc 6360
 10 ttctatattg ttttttttat aagggtggca caaatctttt aaatgagcca aatatctaca 6420
 ttggatttat taaaaactgt ataaattata atttatactc tgaaagggtg tgtgcatctc 6480
 tcttggagaa aatgtataag ttgcaaaca acattaatcc acgttatgta actttttttc 6540
 15 gccggaaagg ccgaaggagg cctgacggag cgtggggctc ctcaccggga gaccgcgcag 6600
 gccccctttt gccgggtcgg ccggggactc agggtgaaat tctaaagctc ctgtatgttg 6660
 aagggtcgcg accgtcgaaa gagcataaga cagggcgat gtatacaggt tcggggcgct 6720
 20 gagaagcgta atacctact cctgtgtttt gggggatctg tgtatgaagg agctacaaag 6780
 tatgagccag cctctccctt gtcttgggtt ccgaatctgg aaaagtcag tccagtcctc 6840
 cctcttaagt gggcaaggct ctccttttat atcttaaggg gataccacat gcaccatctc 6900
 25 cctctttctt gtggagactt acctacctt ttcataaatg gacggagatt tgtatagtig 6960
 ccgtccgaat gacctctga taggacggcc cttacctacc tccacttccg ccgaaagcag 7020
 30 gtgcgacgtg ggattatggc tgtctgtga cgacatgacc agtgtcagac tggtcacaaa 7080
 ttgtctatc ctgtccacca cgcgtcagtt tagcaatcta catgttggcc cttcttcaca 7140
 caacatcttg cctgtaatgg ttaggatgaa gcctggcata tatctaacca ggactaacgt 7200
 35 gccatctcta ggaggtaaca cgctagctcc agctggggac gagcgcttag aagccctctg 7260
 cctgacggga tggggcgagg cgtgcgtcag atcgctgtc gccacctaac ctgcgactct 7320
 accggctctg gactggctac agaccggata aacgagtga ctgcacttcg ttacatgcag 7380
 40 cgtgacacgc tcagccaaac cgcaataaat gtgggttaggt gagccccgct gtgtctacct 7440
 aaccatata cgcggagcaa aaaccacga ggggtcgggg cgctcggcc ctgggggccc 7500
 aggcgggtgc ggtccgaccc cctcgggggg actaagagga gggcgaacac atcacctctg 7560
 45 ggcccagcgt ccccgaggg tgcaggcca cgtggcgat tgtgtctgcc tcaaacctct 7620
 agtcatgata ctccgatcc catgtaccg acagtggccc ccggcgttat gccaggcgga 7680
 50 tgcctctctt taagggaagc ggtcgggcgt gacgccactc ctaaggcctg gtgacagggt 7740
 ggaccggtct ccacaattgg gcagaaaccc aacggtcaca aatcacgcac atcggaatg 7800
 gtaactctac tatcaataat gagcggctc tcaagactg ccacattact cgagtagcac 7860

acgaatcigg acatggcgat tcttttcgtc tggagataig gtaacgtcgc ttigtgtcggc 7920
 gagcglaatt aacgcgcgca cgataigatc taictcgact gccacaaccg cataaccacc 7980
 5 tcatgcgccg caagcgggcg aatgggatta gtggaagcgt gggcgcgaga aacgaggggg 8040
 cgaaatagtg ggcgcgagaa gcgaggagcc gggcacagcg ttggcaagag tataaaggca 8100
 10 ctgaggaaag gatctgttc ctctcttcg ccatcatttc ccttgtcttc gccgcttgcg 8160
 ccctaactcc ttctttccig tgcctactt tgcacacag cgctcgctct caatcttctc 8220
 ttcttccggc gccatggcac ggggtctccg tctgtctgat ggtagcgtgc tgcgccttc 8280
 15 ccgcatcttg agcgagaggg aggtgggct gccgcgccg tcatgccgg aatctgccac 8340
 cggccgggag atagtcacgc tgggcgaggg acgcccggcg ccagactacc cggggcggtc 8400
 cgtcttcttt ctecccttg caatggcagg gctgggtccg ccatcttctt ctctcttcat 8460
 20 ggatgttctg aagtctacg atctccagat ggcgaccctc accccaacg cggatgatgac 8520
 attggccatc ttgcgcacac tgtgcgagat gtcatgtgg glgcgcccat ctcttcggct 8580
 gtccgggtgg ttcttcaccg tgcagtcggt gtgcgccca tggtagttg gtggctgcta 8640
 25 ctccagcca cgggggcccg tgcigaatcg ctacatcccc tgcgccctcc gcaagaagt 8700
 ggacgactgg aagagcgact ggttctacac cccctcgcg gacgaagcgc gcctccgact 8760
 30 tccgagccag ccccgccgc aggcctccag ctggcgggcg ccggtagatc tgggggatgg 8820
 ctatgacgcc gtcttcgacc gccgtggcggg cctacgatcc caggggtca cagggacct 8880
 gggttacggc gactacctcc gtcttcggt tgcgccgtc cagcggcgcg ctggggcgc 8940
 35 ctgggagtac accgggtccg aagactacat gaggaccac cagggagica gatgggacig 9000
 ggctcctgag gattcaaga tagtggtcca acgggtgctg aatctcaact ccatggaggc 9060
 gtccctcatt cccaaggaa tctctctct ctgcagcgat ccagaccgcg cctccatct 9120
 40 gaccattatg acggcggtcg gggcctcaga ggagttagct ccaaagggcc acgacggcgc 9180
 aggcgggagc cgtagggggg atcaatctac ccggggaggg ggtcgtgctt ctgggtctcg 9240
 cgacggaggc ccgaggagca gccgccctgc cgacgcccg gggaagagga agcagggagg 9300
 45 aacacctccc ccatctctc cccgaggggg cggggcgggtg cgtgccagca gcaggcgccc 9360
 ggagggggcc gcgccgacat cgcagcccga gggggagcgc aagaagaagc ggctccgcaa 9420
 50 gatgggggag acagaacct ctacgggaaa cctatcttc cctctaaagt ggtcgtltaa 9480
 ccgaccccc ctgcaggttc tctctaccc atcgtggctg tattcattct ctcaacgcga 9540
 gtcttcact acccatcttg ttctctctt ggtctttct tctgtttcag cgagatccc 9600

tcgcgtccct cccgccattc caagtccggc cagtcagagg ccgaggatcc ggccggccgca 9660
 gagggcccga ggccgggaatc tgaccggcga gagggccgcgg atcgccctacg ggaagccgag 9720
 gagggccccc aggaggccgc ccgggctcgc caggtcgagg aaaccgctcg ggaggaggcc 9780
 gccccggggc gccaggccga ggaagccgct cgggaggagg ccgcccagagc ccaccaggcc 9840
 gaggaagccg ctccgggagaa agccggatit cgccaggacg aggcaatggc gacttccgag 9900
 gcagctcgcg atgaggtcgc gggcgcgctc cttgagccca ctccctcggg cgacgctcag 9960
 gcgacaacti ccggggcagc tggcgacgag gctcgggcg cgctcgcttg gccactccc 10020
 tcaggcgacg cccaggacca accaggctcc agggacatcc ctgagtcgg cacttccatc 10080
 ggccggcccga gccgcgtggc atcctctcca aggcggctct tccccagcc ttctatcgcc 10140
 ccactgagcg cagagcccci tctgcaggcc ttggccggcg caaacaccgc ggtgttgac 10200
 gggcttagtg cccaggctga ggccttgcaa gcagagtggg cggagctcga cgccgcgtgg 10260
 gcgcatgtcg aggagggcg gcgctcagtg gaggcctagg tggaggctgg ccgcaaggca 10320
 caccgccggc atgtctcgga gcttgaagcc cgtagaagg tgttggcgga aatcgccaag 10380
 gaagtggagg aggagcgggg gcctgcccic attgccacca gcgtgaltga cgaggcgag 10440
 gacaccctcc gccttcaata cgggagcttg gaggcggagc tagggaaaaa gctcgacacc 10500
 gccagggggg tgcctgacgc tgcgcgtgcc cgagaacagc gggcggggga gaccgaagcg 10560
 gcgtcccgac ggccggaaga gacctttgag gcgcgcgcca tggcgctgga agagcgcgcc 10620
 tgcgtcgttg agaggatct ggccggaccgc gaggcggccg tcaatctccg ggaggcaaca 10680
 ctggcgggcg acgagtcgc ctgtgccga gaggagtcgc cactccgcct ccacgaggac 10740
 gcgctcaccg agcgggagcg agctctcgag gaggcggagg ccgcggcgca acggctggcg 10800
 gacagccgtg ccttccgga ggcagcgag gaggagcagg cgccggcac tctggaatgt 10860
 gtccgcggcg agaggaccgc actgaaccag caggccgttg acctcgaggc gcgggagaag 10920
 gagctggacg cgaggggcg cagcgacggg gcggctgcgg gcgaaaacga cttagccgcc 10980
 cgccctcgct ctgccgaaca taccatcgcc gatctgcagg gcgcgctaaa ctgctccgcc 11040
 ggggaggctc aggccttccg ctgggcaggc gaggtagggc ccggcatgct ttgggacgcc 11100
 gtctcccgcc tagatcgcg cggtcggcag gtgggctct ggagaggggc gaccgtaaag 11160
 tacgccgcca accatggagg cctcgcccag gcctctcga agatggccag ggctctccaa 11220
 cggctccccg aggagctcga gaagacaatt aagtcacct cgagggacct cgcccaagga 11280
 gcggctggagc tcgtactggc gattaccag gccagggacc ccaattctc tccatggatg 11340

5 gcgcaggatg agtcccccc tgggaccgag gacagcgcg gcgcaggatc gggatgccgc 11400
 cgaccatata gtccacagct tggagggctc agccccctcg ctcgcgttcg cccccaactc 11460
 cgacgaggag gacaatgccg gtggcgcaga cgacagtac gatgaggccg gcgaccggg 11520
 cgtatcgat tgaatcccca agcccccgcc attctttagt tttttcttct tttctttctt 11580
 10 ctaaggccct cggggccctt ttttgtatag atcaacttaa tctgtaata aaaatgaaga 11640
 aatttttgt tcaatttcat ctgtctgtgt gtatgagatg aggatgatct gtgacgttgt 11700
 ccttttgctt cttagcttga ttaagggctc gtgcccagggt cccagtcctc aaaaggcgtg 11760
 15 ggtcggggct agtgcctggg gagatccaca tgcgagacat ggccaggccg ggaacgttgt 11820
 gaccgagggt tatgggtgac ccgattgttg gtttttgccg attccccccc ggagttcacc 11880
 acgccccggg gcacggctcg gtctcgggce ccgtttggcg attttagccg acccgagccc 11940
 20 ccgagggcag gatgagcac gagtgacctt ttcaagta agattctta aaaggaaaaa 12000
 aaaacacaga tacagccctt aggaattga aactgccttt attgaaatac tgaaataaga 12060
 gaaataagaa tgtgcatgtg tggcagcccc cggccaacgc tgcacgcccg aggggggtgcg 12120
 25 ggtttggccc gagcccgaaa cctgacaccc gacccccccc tcaggggtag aagcgacgaa 12180
 ggtgttcgat gtccacggg ttaggcagct caatgccgtc gcccggtggc agccgtatgg 12240
 30 agcccgcccg ggggacgccc accactcgat acggaccctc ccacatttgt gagagcttgc 12300
 tcaatccagc acgcgttttg acgcggcgta ggacgaggtc gtcgacgcag agtgatcggg 12360
 cccggacgtg acgtgatgg tagcgccgca ggctctgtct gtagcgcgcg gctctgaggg 12420
 35 ccgcgcgccg ccttcgctct tccaagtagt cgaggatcct tctgcgaagt tgatcttgat 12480
 cagccctgca gtacatgggt gcccgaggag acctcagggt gagctcggat gggagaaccg 12540
 ctccgcgcc gtagacgagg aagaaaggcg ttccccggt tgcctggctt ggtgtagttc 12600
 40 ggtttgccc gagcaccgt agcaactcct cgatccatga atcgtcgtgc ttcttgaglia 12660
 tgttgaagg ctgtgtttta aggccttga ggatttctga attggcgcg tccacttggc 12720
 cattgtctct ggggtgggca ggtgaggcga agcagagctt gatgccatg tcttcgcagt 12780
 45 agtcgccgaa gagttcacta gtgaattggg tgccattatc cgttaataa cggttaggca 12840
 ctccaaaccg ggccgtgat cctttaatga atttaagtgc ggagtgttta tcatcttga 12900
 50 cgaccggata agcctcgggc cacttagtga acttgtgat cgcgacatac agatactcaa 12960
 acccgcccg gggccgcta aacggctcca ggatacgcag cccctagaca gcaaatggcc 13020
 acgaaagtgg tatggcttgc agggccggg ccggctgat gatgtgttg gcgtggaatt 13080

gacacgcict acatcgccgg accaggctga cgcacatcatt gagagctgtc ggccaataga 13140
 5 aacctggcg aaaagcitta ccaaccaagg tgcgcgaggc ggagtgggct cgcaticgc 13200
 cticattgat atcggcaaga agcacaacgc ctigticccg aggaatgcac ttcaggagga 13260
 ttcatttagc cgcgcgccga tagagggtcc ctctaccag cacglagcgt ttggagatgc 13320
 10 gatggacgcg ttcacitccct tcgcggctct cgggtaaagt cttaictgtg aggtatgctt 13380
 ggaictcggc aatccaagca atcaatctaa gggagctggg agcgcitccc tcgggtccc 13440
 aggcctggac ttcgacgggc ctgcggggcc ggtcaggcgc gtccgtctcc cctaaggggt 13500
 15 cgggtcgcgc cgacggctgg gcaagccitt cticaaaggc gcccggtagg gtctgggctc 13560
 gcgtggacgc gagccgtgag agtctgtcgg caatcatgtt atcccgctg ggcacatgcc 13620
 gaagctcaat cccgtcaaaa tggcgctcca tacgccgtac ttggcgacg taggcgtcca 13680
 20 tctcgggtc agagcaccgg tactccctac agacttgggt aacgaccagc tgggagtcgc 13740
 ctacaccag gaggcggcgg atccccagtc cagctgccac tctagatccg gcaaggagtc 13800
 cctcgtactc tgccatatgt ttagtcgtc gaaagtcgag gcggaccaag tatctgagga 13860
 25 cgtctccgtc cggagaggtc aacgtgacct ccgcaccggc gcccgaaga gacagggagc 13920
 cgtcgaactg cattaccag tggcggtgt gaggcagctg cgaggggtcc gtgtggcct 13980
 30 cggggattga gacgggtcgc ggagccgggg tccactctgc cacaaaatcg gcgagagcct 14040
 ggctcttgat agcgtgacgt ggttcaaagt gcaaatcgaa ctacagaaagt tcgattgccc 14100
 attcaccac ccgtccgtga cctctctgat taigcaagat ttgaccgagg gggtaagacg 14160
 35 taaccacagt gaccgatgc gcctggaaat aatggcgag tttctctgag gccatcagaa 14220
 tagcgtaaag catctctggt gcctgagggt atcgggtttt ggctcccgagg agggcctcac 14280
 taacaaagta gacgggcccgc tgcaccttcc ggtagggccc atccctctcg ctaggggccc 14340
 40 catccctggg gcactcttcg tccaagcagc ctgcgggggc gcactgtct tctgtgtga 14400
 tgacctcggg gtccgaggat aacaggggcg gccttccac agtggctttg gggccgtct 14460
 gggggtcagg ggctcctggc gtctcggac aagcgggcaa agggccaact ccggtcgtca 14520
 45 ggggccctag gcctccgttc ggctcggggg cctcttctcc ctgtcttct ccgggtcag 14580
 tcagcacagg gttagcctcg gggicaaagg gcgatagggt cggccttccc acagtggcct 14640
 cagggccttc ctgggggtcg ggggtctcta gcaccgtctg acaagcgggc agagggcaa 14700
 50 ctccggctgt cgggggcctc agggcacctg tcggctcggg ggctctctt cctgtctct 14760
 tcccgggcca agtcggcaca gggtagggaa gcgcgaaatg agaattatcc tcatcgcgt 14820

ccacaaccaa tgccgcacta actacttgcg gggtcgccgc taagtagagt agcaagggct 14880
 5 cgtctggctc cggggcgacc ataactgggg gagagcttag atacgccttc aactgggtga 14940
 gggcatlitt agcttccctc gtccaggtaa acggtcggga gcgtitgaga agctttaaata 15000
 agggtaacgc ctctctctcc agcctcgata lgaaccgact tagggcggcc atgcaaccgg 15060
 10 tgacgtattg cacatcccta agtttgctgg ggggcgcata cgctctatag cccgtatctt 15120
 ctcggggttg gccitcaatgc cccgggcaga gaccaagaac ccgagaagct tgccgcagg 15180
 tacaccgaac acacacttat cgggggttaa ttttatgcgg gcggagcgga gactctcaaa 15240
 15 agtttccgct agatclatga gtaacgttic ctggttgcgc gtctttacaa ccaagtcata 15300
 gacataagcc tcaatattac gtctaatig gctaccgaaa gaaattcgag tagtacgttg 15360
 aaaagtagga ccigcatlct ttaaccgaa gggcatlgtc gtataacaat aggttccctat 15420
 20 gggggtaatg aacgcagttt ttccctcctc ctccctagcc atgcgaatct gatggtaacc 15480
 agagtatgca tctagaaaac aaaaaaggtc gcaccccgca gtggagtcga caatctgata 15540
 tatgcgaggc agggggtaag gatccttagg acatgccctg ttaaggtcgg tglagtcgat 15600
 25 gcacatccga agcttgccgt tgccttggg aacgaccacc gggttcgcca gccactcggc 15660
 ggggttgacg ctgccatcat attttctggc gatggigggc cggaaccttg gggccaacg 15720
 gacattccga agactcgcca caaaggctct acagccgaca ccaccaaccg ggggcacgga 15780
 30 gggctgattc ccgcgtccgt gttaggtiga cactctggac gaggaagcgc cctccgttgc 15840
 gtgggcagca ctccgtcat taccggcg cgctatgctg gtgcggcggt ccggcccccc 15900
 35 acgcagatct ttctgggtcg aaggagtcga cgaaggagt gcggccgaat ggccaacagc 15960
 ggttgcctct cgtctgccc tccgtcttga cgacgcggag ccggtggtag cagcaccaga 16020
 ggccitgggt gcggaggacc gccaccagc atctaggcgc tgccttgccg tcatgactaa 16080
 40 ttggccacg tcttccagcc atcgttgggc tggagacicc gggtcaggga cgacaggcgg 16140
 gtgacgtaag agcgcgccc cagcttgag cgcgccctgg ggcgtgctgc cgtcgccgta 16200
 gacgaggagg cgacgtccc catctcgcc ttcttctcca tgcgccga tcggtgaagt 16260
 45 cgcggatctt tgcacctct cgagcgctc ccccgctta ggactttggc atggaggag 16320
 cggttggagt cgagctcgac ggcgtgggtt cggctcccc tctcgccac tcacactcgg 16380
 50 agagaggctg tgcgctttg ctgtctggc catcaggctg aacaggaaaa gcttggcgca 16440
 cacggaagag tacgagagct cagaaaaaca cacactgagt cccctacctg gcgcgccaga 16500
 tgacggagcg tggggctcct caccgggaga ccgcgcaggc cccctttgc cggttcggcc 16560

ggggactcaa ggigaaatic taagctctct gtaigtggaa ggtttgcgac cgtcgaaaga 16620
 gcataagaca cgggcgatgt atacagggtc gggccgctga gaagcglaat accctactcc 16680
 5 tgtgttttgg gggatctgtg taigaaggag ctacaaagta tgagccagcc tctcccttgt 16740
 tctgggttcc gaatctggaa aagtcacagc cagtcacagc ccccccctta agtgggcaag 16800
 10 gtcctccctt tatactctaa ggggatacca catgcaccaa ctcctcctt tctgtggaga 16860
 ctaccctat ctcttcataa atggacggag atttgtatag ttgccgtccg aatgaccttc 16920
 tgataggacg gcccatacct accctcacct ccgccgaaag cagggtcgac gtgggattat 16980
 15 ggctgtctgc tgacgacatg accagtgtca gactggtcac aaatgtctca ttccgttcca 17040
 ccacgcgtca gtttagcaat ctacatgttg gcccttcttc acacaacatc ttgccgttaa 17100
 tggttaggat gaagcctggc atatatctaa ccaggactaa cgtgccatct ctaggaggta 17160
 20 acacgctagc tccagctggg gacgagcgcc tagaaacct cgtcctgacg ggaiggggcg 17220
 aggcgtgcgt cagatcgctt gtcgccacct aaccgcgat ctgaccgttc tgtgactggt 17280
 cacagaccgg ataaacgagt gcactgcact tcgttacatg cggcgtgaca cgtcagcca 17340
 25 aaccacaata aatgtggta ggtgagcccc gctgtgtctca cctaaccat acacgcggag 17400
 caaaaacca cgaggggtcg gggcgccctc gccctcgggg ccgaggcggg tgcggtcga 17460
 30 cccctcggg gggactaaga ggagggcgaa cacatcaccc tcgggcccga cgtcccccga 17520
 ggggtgccagg ccacgtgggc gatgtgtctt gccctaaacc tctagtcatg atactcctga 17580
 tcccatgtca ccgacaaggc catccgaatg tattaaggag taaaagttaac aagaaaaaac 17640
 35 accataatgc accaatgtgc atgaccacac accatacact accccaagc acaaaccact 17700
 gagggigaag cctagcacca aacgaccgcc actaagtgtg accaaacgcc gctaggccta 17760
 cggcagcaac acatagatga gacttcgaaa acgatgccac caagggtgtc acgacatcta 17820
 40 ggatgtctgc atcgtccatc taaaaagatg tggttttcac ccagagaaac tcatcaagaa 17880
 ggggagaggg taaccttga cagcgcccca aggaggttac gacgcccga ggcgtagccg 17940
 ctgccgggtc ggtgaaccac cggactaggc ttccgcctag gacctatag ccttgatcgc 18000
 45 agatcaccgt ccaccactca gaaccaccac acagacaaaa ggtagcacgt agcttccacc 18060
 acaccgcacc gacgccccct cgtcggccga ctccatcga ccaccatccc tgagagctgg 18120
 50 cccaggaccc ctccgttcca ccaccgccg gccgccttgc cagtittggc caaaggagaa 18180
 cccgggactg ggtgacattg ctccggcagc ctgagcttcc cccgttggcg agctgtctgc 18240
 tcaatccaac ctagaaactc cccgcaaaag aaggggatga gccttaggaa gggcgagggt 18300

gccgaccggc aacgaggaag acaacccatc gactccagct ccctttgcac taccatctgg 18360
 cccctgcgcca atgccggata cgtctgctgt ccggctccgg cgccaccac ctgcaccccc 18420
 5 tttgcttggc ctccggcgcc ctccctggctg cgtcgcgccc cccagctggc cgctaagggc 18480
 accgcgacgg ccgcccggct accgagggct ggccgcgcca tgggacagct cgcgctggca 18540
 ccagcgagcc acggccgtcg cgcgtgtgcc ggccgagcg agcacaaccg ccagctccaa 18600
 10 gggccgagca tggcactgag ccgcccggcg tggcgcccgg gccggctgca cgtcaccggc 18660
 gcacacgacc gcacgcccgc acgtctcgcc tccgcccgg aggcagcccc atgccattgc 18720
 15 cgcgcacctc gcccgcccgc tggcgagccg ccaccgcgca ccttgcctgag ccgccaccgc 18780
 cgtccctagc cgctctgtgc cggcgccacg ccagatccag gcgcgggatg gccggatccg 18840
 gccctggggg cgccggatcc accgctccc cacaccgcca cggcgctacc accctcgacc 18900
 20 gcagtgaggg ctctctgtgt tggccatcc tcatcgctc gaggaggaag acgccaagaa 18960
 aaaagggctt cgccgctgcc ttccttgttc gctgccggct tggcgcccgg cgagctccgg 19020
 cggcgggcag gtgggggaga agaagtgggg agtgggcagc tagggttttt tcgcccccca 19080
 25 agccgcccgt gcgagagcga cggctggggg gggggggact tccaacctc tccagtggt 19140
 ctagtctcc acgttatgta actcaattg ttttaaccata gaaagtaaga aacctaccag 19200
 30 cgtgttaagc tctctttcat tccctttctt ctctctgggt ttgcttccat cacaatgcaa 19260
 glgaagggtt cttactacc attactccta cacatctaatt tttttctca gatctttcgc 19320
 aggtatata tgaatgctaca ttttatgac ttaagataat ctctttcaca ttacctctg 19380
 35 ctgaaacttt agcttgaacc gtcatttca ccacaattg agcccaattt gcacagagca 19440
 caacgagcaa tagcttgcct ttacgttcat tatttagcat gaactactac taactacca 19500
 agaatcaata caccggttta ataacgcca tttatcagc taatatatgt ttcatccaac 19560
 40 acaccggttt tggcacagtt gcaaacttgc aataaattct ttcctacttc tccatcccat 19620
 aatataacaa attggtatgt ctctcttggc actaagttac tataattatga gatggaggga 19680
 gcacttcttt tcttccaaaa tataagaata tagtatgga ttagatatta tctagattca 19740
 45 cgaattcgat taggttgtct agatttatag ttgtatgtaa tgtataattc ggtaataggt 19800
 tattacctct caggatggag ggagtagttt tgactttttt tttcttataa atcgctttga 19860
 50 tttttatatt agtcaaat tttctgagtt aactaagttt atagaaaaaa attagcaaca 19920
 ttttaagcacc acactagttt caataaat agcatggaat atattttgat aatatattg 19980
 tcttgltta aaaatgctgc tatattttc tataaacgta gtaaaattta aataagttag 20040

actaaaaaaaa atcaaaacga ctataatat gaaatggagg aagtagtaga ctataacaaa 20100
 5 tttaaaccgt gctttgattt tagagcatca ctaataigt agcaataatc tatccctaaa 20160
 atttattttt ttccctaaac tgaanaatagg aagtggaat actcctccat ctaagagaga 20220
 gcctaaattc aataaaaaac taaaaaacta aaggtaggac cctclattaa actaccgcaa 20280
 10 aaaatttatg ttttttttct ctccacgcg cgcagaacag atatctcgat caagttagca 20340
 tgtaaaattt ttaaagagat accittatag actccttccg tatlccaaa agcaaacgga 20400
 tttaaaatct gactcaaata aagatctata tatccaattt acatgacaca tgittcgccg 20460
 15 aatttttata ttaataataa ttaatatitt taaaattaaa ttattagcaa ttgtttgga 20520
 ggatttatca aaacaggatg gacgttggtt ataacagcgt ctagacctag acgcgcitgc 20580
 aaactgcggc caccctttta tcacacaaat ttttgacaat ttgacacttt ccaaaaatta 20640
 20 attttataaa ttaaccgtga ccaaaactta tttaaaaatg atctttttgt tgagcgcaaa 20700
 atcgtatact tcagcgccaa atagcacggc gccgacctcc ccttccccct cccctctatc 20760
 ctccactgct gccgcccacc tctccgtatc agctgcgtcg cgttggtttc cgccggcgct 20820
 25 gctgctgctg caccagtcgg ctagggcggg cgggcatggc gcgcgcgccc gcttcccgcg 20880
 tccgcgccgg cgctgttggc gcccttcgct cggagggtc gaccaaggcg cgagggggcc 20940
 30 gcacgggggg cagtggcgcc gaggacgcac gccacgtgtt cgacgaattg ctccggcgctg 21000
 gcagggggcg ctcgatctac ggcttgaact gcgcccctgc cgacgtcgcg cgtcacagcc 21060
 ccgcggccgc cgtgtccgc tacaaccgca tggcccgagc cggcgcgcgac gaggtaacctc 21120
 35 ccaacttgtg cacttacggc attctcatcg gtctctgctg ctgcgcgggc cgcttggacc 21180
 tgggtttcgc ggcttgggc aatgtcatta agaagggtt tagagtggat gccatgcctt 21240
 tcactccctc gctcaagggc ctctgtgctg acaagaggac gagcgacgca atggacatag 21300
 40 tgcctccgag aatgaccag ctggctgca taccaaatgt ctctccctac aatattcttc 21360
 tcaaggggct gtgtgatgag aacagaagcc aagaagctct cgagctgctc caaatgatgc 21420
 ctgatgatgg aggtgacgc ccacctgatg tgggtcgtat taccactgtc atcaatggct 21480
 45 tcttcaagga gggggatctg gacaaagctt acggtacata ccatgaaatg ctggaccggg 21540
 ggattttacc aaatgttgtt acctacaact ctattattgc tgcgttatgc aaggctcaag 21600
 50 ctatggacaa agccatggag glacttacca gcatgggtta gaatgggtc atgcctaatt 21660
 gcaggacgta taatagtatc gtgcatgggt atgtctctc agggcagccg aaagaggcta 21720
 ttggatttct caaaaagatg cacagtgatg gtgtcgaacc agatgttgtt acttataact 21780

cgctcatgga ttatctttgc aagaacggaa gatgcacgga agctagaaag atgttcgatt 21840
 5 ctatgaccaa gaggggccta aagccctgaaa ttactaccta tggtagccctg cttcaggggt 21900
 atgctaccaa aggagccctt gttgagatgc atggctctctt ggatttgaig gtacgaaacg 21960
 gtatccaccc taatcattat gttttcagca ttctaataig tgcatacgct aaacaaggga 22020
 10 aagtagatca ggcaatgctt gtgttcagca aaatgaggca gcaaggattg aatccggata 22080
 cagtgacctt tggaaacagtt ataggcaatc ttgcaagtc aggcagagta gaagatgcta 22140
 tgcgttattt tgagcagatg atcgatgaaa gactaagccc tggcaacatt gtttataact 22200
 15 cccataatca tagtctctgt atctttgaca aatgggacaa ggctaaagag ttaattcttg 22260
 aaatgttga tgcaggcatc tgtctggaca ctatttctt taattcaata attgacagtc 22320
 attgcaaaga agggaggggt atagaatctg aaaaactctt tgacctgatg gtacgtattg 22380
 20 gtgtgaagcc caatatcatt acgtacagta ctctcatcga tggatattgc ttggcaggta 22440
 agatggatga agcaacgaag ttacttgcca gcatggcttc agttggaatg aaacctgatt 22500
 gtgttacata taatactttg attaatggct actgtaaaaa tagcaggatg gaagatgcgt 22560
 25 tagttctttt tagggagatg gagagcagtg gtgttagtcc tgatattatt acgtataata 22620
 taattctgca aggtttattt caaaccagaa gaactgctgc tgcaaaagaa ctctatgtcg 22680
 ggattaccga aagtggaaacg cagcttgaac tttagcacata caacataatc cttcatgggc 22740
 30 ttgcaaaaaa caatctcact gacgaggcac ttgcaatgtt tcagaacctt tgtttgacgg 22800
 atttacagct ggagactagg acttttaaca ttatgattgg tgcattgctt aaagtggca 22860
 35 gaaatgatga agccaaggat ttgtttgcag ctctctcggc taacggttta gtgccagatg 22920
 ttaggacctt cagtttaatg gcagaaaatc ttatagagca ggggttgcta gaagaattgg 22980
 atgatctatt tctttcaatg gaggagaatg gctgtactgc caactcccgc atgctaaatt 23040
 40 ccattgttag gaaactgta cagaggggtg atataaccag ggctggcact taccgtttca 23100
 tgattgatga gaagcacitc tccctcgaag catccactgc ttcttgttt ttatgtcttt 23160
 tgtctggggg aaaaatatca gaatatcata ggtttctccc tgaaaaatat aagtccttta 23220
 45 tagaatcttt gagctgciga agccttttgc agctttgaaa ttctgtgttg gatttctttt 23280
 ctcttacagt cgtattagag gagggatctt ctctttatgt gtaaatacg aggtatgtat 23340
 50 gtcacctctc cgaattattt ttactctggt tccatagacgg taacaagca attatgttct 23400
 gcccttgaig ccagaaaaaa cacaaaagtt tgcgttatc tctactaacg galcataaag 23460
 gaatttgtaa ctggagtctc aaacttaatt tgtctaggca gtagttttgg cattagatcc 23520

aacatlgigt aggtatcatt tigtigtatc aatclatagg gtllcatlaa atllcgtlla 23580
 tgtgtactgt ttaggtgttg aatagttiga ctgtttttt aactgaacaa aagatactga 23640
 5 aatcgttcca ttcaacaaac aatgtttccg ttaatgaaat tatgttacgt taccitttgt 23700
 ttctttactc acaagigtcc tcttttctta tatcttatag attggtacaa caaatatttg 23760
 attcaatttt gggtttgaac atlgaatgac ctccctgcac tatgtgtgca gctgctcttc 23820
 10 tattcatttt gtgaagtgat gtgagtacct ctcaatccca tccctatgct tctgtgcatg 23880
 ctctattcca attttttacc catatcgatt gtlltctttt atataacagt ccataaagat 23940
 aatcacatca tgacaaagtt atttatttct acagtatagt tatataagta ttaccagatt 24000
 ttccatgaat attttggcat gtgattacaa agaagattat ttgagaaaat ccatgctttt 24060
 atttcatcat ttgttttgaa gtigaacttt aatttatggt gtaaatttca gttattattg 24120
 20 cttagcagctc gtactcttta atggatatac ttactttgtg cttaattctc aatattctcc 24180
 ttcttgttgt tcaggttcaa gaaaatcatt tgttggattc agaacttgtt gtccattttc 24240
 ttcttaaat attaaatctt ccagtgatc ttgttgattc caaagcacca tcataggtt 24300
 25 ccaaacttct tggaaatcagt aaagtcaaaa tgcctaatgg atcaataaag gattctgact 24360
 gcatttcaga ggaaatcttt tcaaaagttg aagagattct cttaagctgt caagtatca 24420
 agtcgctcga caaagatgac aagaaaacaa caaggccaga acigtgtcca aagtggcttg 24480
 30 ctltgttgac aatggaaaat gcatgcttgt ctgctgtttc agtagagggt aagtittaat 24540
 caaatttctt ggcatgatt tccctttatg accattatat ttattatat gagccaaata 24600
 35 agcagttgtc aactgtcat aagttacata gcacctattt gcaatattca tgggttgttt 24660
 gcttagccct ttctttcacc tgcctttgat tgaatgattc catctgtgtt gcagaattga 24720
 attggtagtag tggactgcac tagaagcacc tatggccatt gtcatactag gaaggttttc 24780
 40 ccttatcaaa tatttgattg ttacagagac ttctgacaca gtgtccagag ttggaggaaa 24840
 ttttaaagag acattaaagg agatgggagg tcttgatagt atttttgacg ttatggtgga 24900
 ttttcatcca acattggaga tgagaatcgc ctacatcgc atattttaca ttcttttgt 24960
 45 tcaactctaa tagattgtgc aggttgttc ctlltcgcca tttagcttt aatgcgcttg 25020
 aagccacatg aaagtaatgc ttgtccagat acatagccaa aggttgttat attttggggc 25080
 atggaaaaatg cttaggttag taactatttt catcaggaca tggaaaatg gcigcaacac 25140
 50 aaattatgtt gtlltaigt tcaaaaaatg ttttttaata cttttttatt ctgcatgtgg 25200
 tgttagtatc ttacagttcc tctgatgatt atatcccca cgataataac acttgaaacg 25260

alaataacac ttgacataac tacaccaagt gaacattatt catttggatg ttacttttcc 25320
 agctatactt gctgttcttg catgtgtaag caagtittga gtaaatigcg cattaatlla 25380
 5 aalgcttggg gticctatct gtgtactttt latccccaa ctaataatgc aatcatatla 25440
 cgctgataaa ctgaataaai aaattaacaa tatacttcig gtggcaaacc ttgtgtatca 25500
 10 gaatcicata aaggatacat ccacttcagc ttiggaccga aatgaaggaa catctttgca 25560
 aagtgctgct ctctcttga aalgttgaa aatatggaa aatgccalat ttctaagcga 25620
 tgataacaag gtaatgctcc ttatatgtic tgtttcagtt tagtacctat ttcttcttic 25680
 15 tgtactatct tctctctga ttgttctgt gcaaaalggt caaacagtgc gactttgtat 25740
 gtctgtctaa caatttctt ttcttctga aaaagcaata tgaactctta cattcatttt 25800
 gcttcttgca gaccatttg ctaatatga gtgaaaaat gaaccgaaa cgctcttgc 25860
 20 tttcttttgt tgggticatt atcaatacia ttgagttatt atcaggtatt ttcttaata 25920
 atacaatgtg ttegetaaca caataaaaig ttttaaacat ccagtaigt aaagtgcag 25980
 tctgacgctt atttgtttt gctgcagctc ttcaataact tcagaattct tctgttgttt 26040
 25 ccagctctac atatccgaaa tgcctaaag tctctcaaca gagttactct gglaataaca 26100
 aacaccaatt ttgtttgatc agttgatctc gtggccttt ctatgcactg tctcaatata 26160
 30 gtttggctgc cattcaagtc tcaactacaga tgttgaactt ggcctgacac caaatatlla 26220
 taaaatgcta ccgatattt ttaatatctc atgtttcctg acccagatta tctgttggg 26280
 tctcgtata agtttaatta gtgacattct tgaagctttg ttatgcagca gatgtcatgg 26340
 35 ggggaacttc atttaatgat ggaaagagca agaactcgaa aaaaaaaaaac tttgtcgaa 26400
 ccagacagct cattgttgc tcttcaaaa atcagaagtt tctcatatta ctatatcttc 26460
 tggtagtgat gctgggtctg cacagaaggc attcaattgt tctccattla tatcaagcaa 26520
 40 tggggcatca agtggttcat taggcgagag gcacagcaat ggtagtggtt tgaagttgaa 26580
 tataaaaaag gatcgtggca atgcaaatcc aattagaggc tcaactggat ggatttcaat 26640
 aagagcgcac agtctgatg ggaactccag agaaatggca aaaagactcc gtctatctla 26700
 45 aaaigtatc accgacagt gtgggtgga tgacctttt gcatttgacc gccgctcgg 26760
 cgctgccacc acgtaatgc ccacgtcgt gccccgcgt ccacgtctc gaccgcgac 26820
 50 gglaatcaca cgcatctga ggccgccgt agctgatac ttctcatccg gtigatttgt 26880
 gattttggcg tttttgcagt ggtatggcg gggggcgacc gtggccgagg cgtggagtgc 26940
 catccgcatc aggggtatc ggccgcgtc ctccgccctg gtccgcaggc ttggcgggcg 27000

agctggcggc ggagggagac tglgtgaga tggatttcg ccgctggcgg tgcctacc 27060
 atgggggatt cgcgcaggc gctctcaggt ttgcagcctc cccactctc tcccttttt 27120
 5 ttttttttt tctcgcaaaa tglgtgtga tgltcgtctc gctggcctgg cctcatagcc 27180
 attaatgtag ttgtcggaa cattacatt tggacgttg ttggcaattg ctttcaaaa 27240
 10 tgtggaattg tggaggggag aaaaatcatt tgaacctgca gtgacaaaat tggcatctct 27300
 aattttaaaa ctgaagggtt ggaaatcaaa cataatcatt gccagcgcat cattcttgtt 27360
 aaccaccatg atatatgtt ggtataaca gttagctcca caccaacctt gaagggtgca 27420
 15 atagaatgtt tagtataaat tgaggagaac aggcagttgt taagacttc taaagaactt 27480
 gtagcagcta atactagcta ttgtgcatit gtgttcatg gaatttgagc agcaatggat 27540
 atttcttact aagaatgtat algcaaaaca aaaaactatg tctatacagt ttacatgtaa 27600
 20 tglcggatg caaataaaa catgtacatg gacaaactca tgggattcat accgaattcc 27660
 agaattgcat ttcttatgtg gttacttttg ttgttgatit ggttaccaga catcgatgtg 27720
 atttcaaggg tcagaggggt ttgttctac gcggcggcgt cagttgcagc aatctttttg 27780
 25 tttgtcgcca tggttgtgtt tcatccactt gtgcctctat ttgaccgata ccggaggaga 27840
 gttcaggaaa aaaatttgaa aataccatt ttttgaaaaa gatttacgtt tatatacact 27900
 30 agtatgaaga atttgcgaaa atataactaa tccgcagatc ggttatgcgg gagcgcaaca 27960
 aaagtatggc gtggcggcgc ggagtggacg gccgagggct tcgcgcggaa tggggctgcg 28020
 ggaccgagcc agtctgcctt gccggtaacg cggaaccggt acgtcccgcc agcgccagtg 28080
 35 tgcggaaccg cggcgccaac atttttttac tgcattggcac tgtgtttaat actgtttgac 28140
 actgtttctg gtactgtttt acacagttcc cgggtcagtt ccgcacaatg gagcgcgccg 28200
 accgaccatg aacaatgtgt gaacagtgtt gcacagggtt aaaacagtgt ataaactgcg 28260
 40 ctgcacagtg ctggagtcgc tggccactgc ggttccgcgt tttggaaccg cgggaccgtc 28320
 gcatccgc gttttggagc tggcgacca tgacggttcc gcgcaggatc gtcggtcccg 28380
 tattttgaat ctgcggaacc gtgcgttcc cgcgttcca ttctcgggga tgcgtatatt 28440
 45 ttataaaaac ctctccatgc atgtatata acataaatla ttgaaaaaat aagtatatit 28500
 gcaaatitit ttcgagagct cagcactaca ttgcaaagat ttgggcaact ctgacaatit 28560
 50 ccatgttcta caagcttgac gtcgagggaa tggagaacct gccaccgaat agtagccctg 28620
 ctatctatgt tgcgaacct cagagttttt tggatatcta tacccttcta actctaggaa 28680
 ggtgtttcaa gtttataagc aagacaagta tatttatgtt ccgaattatt tgatgggcaa 28740

tgtatctctt aggaglaatt cctttgcggc gtaaggacag caggagccag ctggtaaggc 28800
 tgtagctca tccctgcttt ctttaagtaga cataatgca attacagaat ttggtaaaca 28860
 5 aacaagattt tatgaatcat atatgatttt ggggaaaaca ccaaactctc ttgggtggct 28920
 gccttgaaca tagtictatt cacacagita tagcaccitc tttaaaatga agaactttgt 28980
 10 tgcatacaca taiggccaaa ccacataatg aattttgttt atttctatct ttgaatgita 29040
 gcaccttatt ttcaigcata tcatgctaatt ttgcttgccc acgttgagtg ggaatttttt 29100
 tccatgtttt ataatttata tatgttctag acttctagtc cacaatttat ctacitcatg 29160
 15 ttcttgagcc tctagiatgg ctggtagcag actagggtgt gagtgctgtc catttttgca 29220
 gactgaagag aggagaaata caggactgtc cgttgttagt cagatttgta aaaatagact 29280
 ctgatgtagt ttattttagc cctattttta tatttaacaa tacaatatata taacgtatcc 29340
 20 taagaactta tcgtaattta ggagaagttg ctctgttcat taaattaaac tgtgaagtaa 29400
 aaatgtgtgc tcgagctgti caatgcaatc ctgtgttctt gtttgaagat atgggttagg 29460
 gcaggctagg atcgaacact gaatggtaag actgcttctg ctttcaattg tgcacttgg 29520
 25 gctgccacgc cgattaaagca gtagaacaaa glaattttgt cgtgcacaaa tgagtatat 29580
 ttcatigaaa atcgaagtga aaatgaacca aaagatagaa gaaaagggga aacttggtaa 29640
 ttatatctc cacaatttta ttggtaagat ttgatattag acgctcgatt acttggctta 29700
 30 agttaaggat atcaaatttg gggaagcacc aaaggaatta ttgtgaagga gttgtgggtg 29760
 cataacgita tctactagtt caaatccctag tgactatgaa tattaatgag taaggtaagg 29820
 35 gatttatgti taattttagt ttctttaaga ttgtgtccga gtacaccatt cggtaagtgt 29880
 aataatgttt tgtatggat tcacttgtgt tacgtgcatg tgcttttacc ttttcaattg 29940
 ttctgtcggt ctgggtatga atttgacgag attcatggt cagctcaaca tatcagttac 30000
 40 tgcgtgtcaa gcgatcttat atggtaigcg cacaagcgt tglatacgga tatgacagta 30060
 taatgtgtga tattgatacg atgttccctt cctttataaa ggaacaaaga ctttttttaa 30120
 aaaaagaagg ggtattacta aaaacaaaaa tgtcaaaaac aaaatatcag tgcacatggc 30180
 45 aagttgtcac gagcaatagc ttgcccctac gttcattatt tagcatgtac tactactaac 30240
 tacgcaaaaa tcaattcacc gattattaaa ctgttaacat catttttagca cgttaacata 30300
 50 tgtttcattc aacacaccgg ttttggcaca ttacaaact tgcaaagttg caatactccc 30360
 ttcgttacat agcataagag attttagggt aatgtgacac atctatccaa attcattata 30420
 ctagaatgta tcaccgctc cagccggga gggagagcgc cgccgggtga gaaaggggga 30480

gggagiggic gaggggaacc agtagggigc cctccccgic gccgcctccc cgtggccgcg 30540
 ccggcgagac aggaggaaga gggggagatg gagcggcgcc gccggtaggg gcgcgcgtgc 30600
 gcggggggggg ggggggggga gcggcgacgc cggtagggaa gggaagggga glggtaggctt 30660
 tgagagagat aggggagagg gaaaatgatt ttagagttag gglttgggct gctgagtttt 30720
 tatatagatc gggatcaatc aggaccgtcc atcagatcgg acaactacgg ttctctccgc 30780
 gttgggcccgg gtagccactcc taggttggcc acactatigg gccacatgia cgctccgcgt 30840
 gaaataagtt cactttaggt cctttaagtt gccctcgaat lgttcccagg ccggccgcac 30900
 tattgggcca ccccataggc catgtgtacg ctccgcacag aataatttcg ctttagctcc 30960
 ctttaatttgt cccctcaaac ttctaaaacc agtgcgaatc ttttaatttt agttcaccca 31020
 ttgcaactca cgggcatatt tgctagtac atataataatg aaacgaagga ttagtagac 31080
 tatagaattt aaactgtgtc ttcatttttag agcatcacia acigtatttt agatttttat 31140
 ttaaaataat gcagaaatga tgtttttatt atgaaaatta gcaataaagc tcccaaaatt 31200
 tcaaaaaaaaa attaaaagag atttattaat catggttaat ttaattaaaa attaaatcia 31260
 accatatcat attatttcac ggtaggtgat gaggaaatgg cagctgctat cacttatgg 31320
 gggagagaag gggcattgtt tatttttata actatctctt ataactccca tgaactata 31380
 aaataaatat aatcattatc ataacattag ttttttcca ttgcaacgca agggtaattt 31440
 ttcagtacaa taaaaaaaaa aaagtgggcc attctgaacg gaaatttctg gtttttttct 31500
 ccaagagcgc cgcacacaac tgcgcaagag atcgaatcgc ataccctgc tctctccga 31560
 tctcctacac catccctgcc atctcttcc cctccactgg ctgctgctgc acctgtcagc 31620
 tagggcgggc atggcgcgcc gcgcgccttc ccgcgcgtct ggccgcttc gctcggaggg 31680
 ctcgatccaa gggcgagggg gccgcgcggg gggcagtggc ggtggcgcgg aggacgcacg 31740
 ccacgtgttc gacgaattgc tccgtcgtgg cataccagat gtcttctctt acaatatctt 31800
 tctcaacggg ctgtgtgatg agaacagaag ccaagaagct ctgagctac tgcacataat 31860
 ggctgatgat ggaggtagct gccacctga tgggtgtcgt tacagcacgg tcatcaatgg 31920
 ctctctcaag gagggggatc tggacaaaac ttacagtaca tacaatgaaa tgcttagacca 31980
 gaggatttcg ccaaatgttg tgacctaca cttattatt gctgcgctat gcaaggctca 32040
 aactgtggac aaggccaagg aggtacttac caccatgggt aagagtgtg tcatgcciga 32100
 ttgcatgaca tataatagia ttgtgcatgg gtttgcctt tcagggcagc cgaaagaggc 32160
 tattgtattt ctcaaaaaga tgcgcagta tgggtcga ccagatgtt ttacttataa 32220

ctgcctcatg gattatcttt gcaagaacgg aagaatgcacg gaagcaagaa agatttttga 32280
 ttctatgacc aagagggggcc taaagcciga aattactacc tatgggtacc lgcttcaggg 32340
 5 glatgctacc aaaggagccc ttgttgagat gcatggctc ttggatttga tggtagaaa 32400
 cggtagccac cctaataat atgttttcag cactctagta tgtgcatag ctaaacaaga 32460
 10 gaaagtagaa gaggcaatgc ttgtgttcag caaatgagg cagcaaggat tgaatccgaa 32520
 tgcagtgacg tatggagcag ttataggcat acttgcag lcaggcagag tagaagatgc 32580
 tatgctttat tttaggcaga tgaatgatga aggactaagc cctggcaaca ttgtttataa 32640
 15 ctccctaatt catggtttgt gcacctgtaa caaatgggag agagctgaag agttaattct 32700
 tgaatgttg gatcgaggca tctgtctgaa cactatttc tttaattcaa taattgacag 32760
 tcattgcaaa gaaggagggg ttatagaatc tgaanaactc tttagacctga tggtagtat 32820
 20 tgggttgaag cccgatatca ttacgtacag tactctcacc gatggatatt gcttggcagg 32880
 taagatggat gaagcaacga agttacttgc cagcatggtc tcagttagaa tgaacctga 32940
 ttgtgttaca tatagttact ttgattaatgg ctactgtaaa attagcagga tgaagatgc 33000
 25 gttagtcttt tttagggaga tggagagcag tgggtttagt cctgatatta ttacgtataa 33060
 tataattctg caaggtttat ttcaaacagc aagaactgct gctgcaaaag aactctatgt 33120
 30 cgggattacc aaaagtggaa ggcagcttga acttagcaca tacaacataa tccctcatgg 33180
 actttgcaaa aacaaactca ctgatgatgc acttcggatg tticagaacc tatgtttgat 33240
 ggatttgaag cttagaggca ggactttcaa cattatgatt gatgcattgc ttaaagttag 33300
 35 cagaaatgat gaagccaagg atttgtttgt tgccttctcg tctaacggtt tagtgccgaa 33360
 ttattggacg tacaggttga tggctgaaaa tattatagga caggggttgc tagaagaatt 33420
 ggatcaactc ttcttttcaa tggaggacaa tggctgtact gttgactctg gcatgclaaa 33480
 40 ttctattgtt aggaactgt tgcagagagg tgagataacc agggctggca cttaccttc 33540
 catgatgat gagaagcact ttccctcga agcatccact gcttccctgt ttatagatct 33600
 ttgtctggg ggaaaatatc aagaatatca tagatttctc ccgaaaaat acaagtcctt 33660
 45 tatagaatct ttgagctgt gaagcatttt gcagcttga aattctgtgt tggaaattct 33720
 ttctcctaca gtccgattag aggagggatc ttctctgtat gtgtaaatag cgaggtatgt 33780
 atgtaccctc tccgaattat ttgactgtg gtccctggac tglaaacaag ctattatctt 33840
 50 clggtgttga tgccagaaaa aacacaaaag ttgtctgta tctctactaa cggatcataa 33900
 aggggtttgt aactggagtt tcaaacitaa ggtatctagg cagtaggtat atattgatcc 33960

5 tacatcttat gatcttaaga tgataicctt ctcatatcc tctgctgaaa ctttagcttg 34020
 aaccgctatc tacaccacaa tttagagcccc ttagcacaga gcacaacgag caatagcttg 34080
 10 ccccttacgtt cattatittag catgcactac tactaactac ccaataatca atacatcggt 34140
 tattaaactg ttgttacagt ttaataaigt catttitaica cgttaacata tgtttcattc 34200
 aacaccacac cgglltllggc acagtllgcaa actllgcaata acatttllac tacttctccg 34260
 15 ccccataata taacaatctc gttccatact atattlgtat attacaggat ggatgaagia 34320
 ctcttttctt tccaaaatat aagaatctag tactagatta gatattaitt ggattcacga 34380
 attlgattag gctgtctaga ttgttagtcg tatgtaatgt ctaattcggt aataggttat 34440
 tacttctttg gatggaggga gtagtttita ttctgtactc cctccgtttc atattataag 34500
 ttgttttgac ttttttctta gtcaaatitt attagatttg attaaattta tagaaaaaaa 34560
 20 ttagcaacat ttaagcacca cattagtttc attaaatgta gcatggaata tatttttata 34620
 atatgtttgt tttttattaa aatgclacta tatttttcta taaatgtagt caaatittaaa 34680
 gaagtltgat tatgaaaaaa tcaaaatgac atataataig aaactgagga tglagcagac 34740
 25 tatagcaaat ttaaactatg cttttatitt agagcatcac caaaagattt gcaataattt 34800
 atccctaaaa ttcaagtttt gggtttctta aactgaaaaat aggaagtiga aaatcttttc 34860
 30 cgtccaagag atagcclaaa tcttaictta actaatlaaa atattcataa ttttcttttc 34920
 gtcacattaa attttcgtcc glaaaictga ttgaaatcca atlggacaat ccaaaaaata 34980
 gagaanaaga acagaaaaaa taataaaaag cacacaaatc ttatctcaat cccgcgggaa 35040
 35 gctgccgacg ccgccgaatc cgctcgagcg ccgccgccgc cgctcacggg gaacgatgtc 35100
 gctgcgtcgc cacgcgggat gggagggcgc cgctgccact gcttgggaga taggatatgg 35160
 agagagaagg aatgtigagg gttagggta ggtttttccc cgtccgtatc tlcagcgaca 35220
 40 cggaggcgat ccaagctgtc catcagatcg gacggctcag aatgccacca tgcgcgggcc 35280
 gcgcatgctt gatgggccga gggaaggccg gagggctgaa caaacgcaat caaaggagga 35340
 gtgggaggag glaaattaga atttatttgc gggctgagat aglaaatgga ctgaaaatgg 35400
 45 cccatagaga aattgggaat ttatittaaa taaatgtiga aaagggtgtt atattatcaa 35460
 aattaaaaat taagctccga aaattctaaa aaatattcaa agagcattat taatcatggt 35520
 50 taatttaata aaaattaaa ccaaccatai catattattt cacggcgcgc ggtaggaaaa 35580
 tgcgcagctg ttgtcgttta cgttgggaga gaaggacat tgtttatttc cagaactatc 35640
 ttttataact cccatggaac tttaaaaata atataatcat tattatagca ttagtitttt 35700

1 tctgtctttt ttttccccaa gagegcgcgcg cagaagagat cgatcgcgat ctccctgccc 35760
 5 cgacgtcgcc ggccgatctc tcattctctc cagccctgc tcgtcgccga tctcttacac 35820
 catccctgcc atctctctct tcccccccc tctatctctc actggcgccg cccacctctc 35880
 cgtataagac aaactgcgtt gcggcgctgg tttccgcgg cgctgctgct gcacctgtca 35940
 10 gctagggcag gcatggcgcg ccgcgcgcgt tcccgcgctg ttggcgccct tcgctcgga 36000
 ggctcgatcc aagggcgagg agggcgcgcg gggggcagtg gcgcgagga cgcacgccac 36060
 gtgttcgagg aatgtctccg gcgtggcagg ggcgccctga tctacggctt gaaccgcgcc 36120
 15 ctgcgcgacg tcgcgcgtca cagccccgcg gccgcgctgt cccgctacaa ccgcattggc 36180
 cgagccggcg ccggcaaggt aactcccacc gtgcacacct atggcattct catcggttgc 36240
 tgcgtccgcg cgggcgcgtt ggacctcggt ttgcggccct tgggcaatgt cgtcaagaag 36300
 20 ggatttagag tgggaagcat caccttacct cctctgtctc agggccctcg tccgacaag 36360
 aggacgagcg acgcaatgga catagtctc cgcagaatga ccgagctcag ctgcattgcca 36420
 gatgtttct cctgcacat tcttctcaag ggtctgtgtg atgagaacag aagccaagaa 36480
 25 gctctcgagc tgcgtcacat gatggctgat gatcgaggag gaggtagcgc acctgatgtg 36540
 gtgtcgtata ccactgtcat caatggcttc ttcaaagagg gggattcaga caaagcttac 36600
 agtacatacc algaatgct tgcattggag atttcaccag atgttgtagc ttacagctct 36660
 30 attattgtcg cgttatgcaa gggtaagct atggacaaag ccatggaggt acctaccacg 36720
 atggttaaga atgggtgcat gcctaattgc atgacatata atagtattct gcatggatat 36780
 35 tgcctttcag agcagccgaa agaggctatt ggatttctca aaaagatgcg cagtatgggt 36840
 gtcgaaccag atgttggtac ttataactcg ctcatggatt atctttgcaa gaacggaaga 36900
 tccaccgaag ctagaaagat tttagattct atgaccaaga ggggcctaga gcctgatatt 36960
 40 gctacctatt glacctgct tcaggggtat gctaccaaag gagcccttgt tgagatgcat 37020
 gctctcttgg atttgatgg acgaaacggc atccacctcg atcatcatgt attcaacatt 37080
 ctaatatgtg catacgctaa acaagagaaa gtagatgagg caatgcttgt attcagcaaa 37140
 45 atgaggcagc atggattgaa tccgaatgia gtgacgtatg gagcagttat aggcatatt 37200
 tgcaagtcag gcagtgtaga cgatgctatg ctttattttg agcagatgat cgatgaagga 37260
 50 ctaaccctta acattatgt gtataccctc ctaattcata gtctctgtat cttagacaaa 37320
 tgggacaagg ctgaagagtt aattcttgaa atgttggatc gaggcattcg tctgaacact 37380
 attttcttta attcaataat tcacagtcac tgcaaagaag ggagggttat agaattcgaa 37440

aaactctttg accigatggt acgiattggt gtgaagccca atgtcattac gtacagtact 37500
 ctcacgatg gatattgctt ggcaggtaag atggatgaag caacgaagtt actciccagc 37560
 5 atgttctcag ttggaatgaa accigattgt gtacacata atactttgat laalggctac 37620
 tgiagagtta gcaggatgga tgacgcatta gcctttttca aagagatggt gagcagtggt 37680
 gttagtctta atattattac gtataacata attctgcaag gtttatttca taccagaaga 37740
 10 actgctgctg caaaagaact ctatgtcggg attaccaaaa gtggaacgca gcttgaactt 37800
 agcacataca acataaactt tcatgggctt tgcataaaca alctcacatga cgaggcactt 37860
 15 cgaatgtttc agaacctatg ttgacggat ttacagctgg agactaggac tttaacattt 37920
 atgattgggt cattgcttaa agttggcaga atgatgaag ccaaggattt gtttgcagct 37980
 ctctcggcta acggtttagt gccagatgtt aggacctaca gtttaattggc agaaaatctt 38040
 20 atagagcagg ggttgctaga agaattggat gatctatttc tttaaatgga ggagaatggc 38100
 tgiactgcca actcccgcat gctaaattcc attgtagga aactgttaca gaggggtgat 38160
 ataaccaggg ctggcacitc cctttccatg atgatgaga agcacttttc cctcgaagca 38220
 25 tccactgctt ccttggtata gatctttgt ctgggggaaa atatcaagaa taicatagat 38280
 ttctccctga aaaatacaag tccittatag aatctttgag ctgctgaagc attttgcagc 38340
 ttigaaattc tgggttgaa ttcttttctc ctacagtcgg attagaggag ggatcttctc 38400
 30 tgtatgtgta aatagcgagg tatgtatgtc acctctccga attattttga ctgtggttcc 38460
 tggactgtaa acaagctatt atctcttggt gtgatgcca gaaaaaacac aaaagtittgt 38520
 35 cgittatctt actaacggat cataaagggg ttgttaactg gagtttcaa cttlaaggtat 38580
 ctaggcagta gttttagacat tagatccaac attggtatgt attcatttgt gtgtatcaat 38640
 ctatagggtt tcattaaatt tcaatttgtt acigttagg tgttgaatat attgttttac 38700
 40 ttgtttttta actgaacaaa agatagctga agctttgttc ttaccaaaat gcagtagtga 38760
 tcatcacaat atattttttt acggaacagg agatgtata aaatggtttc catcggcggc 38820
 caacggcgac cgtctgctc tgaccacca cccaalccat ccatccactc gccgccgcc 38880
 45 ctgatccaag cctccgccgc gcgacagcga cgcaccgccg tgcagaggag gaggcgtgag 38940
 ccccatgggg accctcttcc ggccgcgtaa tgcgcgtgca cggtaaccac gcgcctctcg 39000
 aggcctccgc cgctagctga tctcttctca tccgtttgg gtttgggttt gtgatttggg 39060
 50 tgttttttcc gcagcgggtg tgggtgtgtt ggttgcggcg ggagggggcg gtggccgcgg 39120
 ccgtggcgig gagtgccagc tgcacgggtt gcaccgccgc cgggttccgc aggttgttgt 39180

ggcgacggcg agctgaggag gcgaggagg agtggtagg gacacaggca ggcaggctct 39240
 caaggctaag ctgtttacag gtactgagac tagttactaa ttactttgat aatcagata 39300
 5 aataagcttg tglagtgtaa tggcattgtg cattctgca ctgttaaatt ttacagaaga 39360
 tggtcattca atttgaacct gcalctaata ttttagtggg ttgagtttat tctcccagtc 39420
 acagagtga agaggcaagt aacctgttag agaggactga acattaacac ctctgtgtcg 39480
 10 attaaaaatg accaaagagc atcaaactg taticgaggc tgttacttta atatggccca 39540
 ttaatttgtt tagttggcta tgtacatcct agttgggtga gtgtgtgtga aaacggaata 39600
 15 cgggtgtcgg atggacgagg tgcctcaag cgattaatcg taatacggat gattaaacgg 39660
 aattatatgg atttttggcg ttgcactaa gatgtacata attgatgtta atggcaatgg 39720
 tggagacaaa atgcatcctc ttaataaaaa atatttgtat aaatctctaa ctatattatg 39780
 20 aaaatgccat ttattagttc aatagatctc aacactgatg gttagtagcg caatagcatt 39840
 gggcttgtta gtcaaaatag tgcagctggg ctgcaagtgt caagtttatg ttagtttcat 39900
 aaacagacat ctgattgtc gataaataac cgactaatcg tgccatacaa ctgtataatt 39960
 25 actctgaaat agtaatgttg ctccgacttg atgatacggg acggtctggc taccgtttcc 40020
 gttttgacag acgattaaac ggcgtgccc gtgcacttcc acaacactga gttggtgtaa 40080
 atgccagtta ccatttctat gatctaaaaa aatcaactct tttagtatat ttcaaaaac 40140
 30 gaaaattcag tacacatgca tgaatcttaa tcttcatact tagctcgta caaaatcaac 40200
 aaaggcaccg tgcagctgg tgcacattag ctagtctgta cttagcatta tccactagca 40260
 35 ccttattttc atgataatca tgciaatttg ctgcccacg ttgagtggga atttttttcc 40320
 atgttttata atttatatat gtcttagact tctacttcat gtctctgagc ctctagtatg 40380
 gctggttagca gactagggtc tgaatgcgtt ccttttttgc agactgaaga gaggagaaat 40440
 40 acaagactgt ccgttgttag tcagattgtt aaaaatagac actgatgtag ttattttttg 40500
 cccctatttt atatttaaca atacaatat ataacgtatc ctagaattt atcgtaattt 40560
 aggagaagtt gctcgtttca ttaaaftaaa ttgggaagta aaaatgtgtg ctcgagtatg 40620
 45 tcaatgcaat cctgtgttct tgttgaaga tatggtgtag ggcaggccag gattgaacac 40680
 tgaatggtaa gactgcttct gctttcagac gttattgtca aatttttagc tagttgcaat 40740
 tagtgcgtc acgccattta agcaglagaa caaagtaatt ttgtcgtgac aaatgagtta 40800
 50 tatctcttg aaaatcgaag cgaaaacgaa ccaaagata gaagaaaagg gaaacttggg 40860
 aattactcca caaagagaac aaatttatgt gtaagatttg atatgagatg ctcgattact 40920

1 tggcttaagt taacaatata aaatttgggg aagcaccaaa agaattatig tgacttaagt 40980
 2 taaagataatc aaatttgggg aagcaccaaa ggaattatig tgatggagtt gtgggtgcat 41040
 5 aacgttatit gctttgttca aatcctagtg actatgaata tgaatattaa tgcgttaaggt 41100
 aaggaattta ttgttaattt taggttcttt acgatttgtt cgggggacgc cattcggtaa 41160
 10 ctgtaataat gttttgtatt ggattcactt gttttacatg cacgcactaa acatgtgctt 41220
 taccctttca ttgtttgtg cgttcgtcgt ttgaatttga cgagattcca tggtcagctc 41280
 aacatgtcag ttactgcgtg tcaagcagtt actgcgtgtc aagcgatctt atatggtatg 41340
 15 cgcacaagcg attgtatagc gatafgacag tataacgtgt gataatgatt tttttatata 41400
 aaaaaatagc atgttacttt ccttcataaa ggaacaaaga cttttttttt aaaaaaaga 41460
 aggggtatta ctaaaaacaa aaatgtcaaa aacaaaat atcagtcacat ggcaagtgtg 41520
 20 ctcggtcaatt ttttgtctgt actttaacaa aaaatatttc tatatggtat tttttacaag 41580
 ggtgtcacaa atatttttaa tttagccaaac atctgcattt tattaaaaac tgtataaatt 41640
 ataatttata ctctaaaagg ttgtgtacat ctctcttgga gaaaatgtat aagtltcgaa 41700
 25 caaacattaa tccacgttat ataagtcatt ctgtattta accatagaaa gtaagaaacc 41760
 tactagcgtg ttaagctaag ctctctttca ttctctttct tcttcttggt ttgtcttcaa 41820
 30 tcaattgtca agtgaagggt tcttaactac cattactctt actcaccaaa tttttttctc 41880
 agatctttcg taggtatata ttgatcctac atcttatgat ctttaagatga tacccttctc 41940
 attatctctt gctgaaactt tagcttgaac cgtcatctac accacaattt gagcccctta 42000
 35 gcacagagca caacgagcaa tagcttgccc ttacgttcat tatttagcat gcactactac 42060
 taactacceca ataatacaata catcggttat taaactgttt gtacagttaa ataatgtcat 42120
 tttatcacgt taacatatgt ttcatcaaac accacaccgg ttttggcaca gttgcaaact 42180
 40 tgcaataaca tttttactac ttctccaccc cataatataa caatctcggt ccatactaga 42240
 ttgtatattt acgggacgga tgaagtaatt ctttcttcc aaaatataag aatatagtac 42300
 tagattagat attatttggga ttacacgaatt tgattaggct atctagattt gtagtcgtac 42360
 45 glaattgtca attcggtaat aggtatttac ctcttggat ggaggaggia gtttttattt 42420
 cgtactccct cgtttcata ttataagttg ttgtacttt ttctttagtc aaattttatt 42480
 50 gagtttgact aaatttatag aaaaaaatta gcaacattta agcaccacat tagtttcatt 42540
 aaatgtagca ttgaatatat ttttataata ttgtttttt tttattaaaa tgctactata 42600
 tttttctata aatgtagcca aatttaaaga agtttgatta cgaaaaaaaa tcaaaatgac 42660

atataataig aaactgagga tgtagcagac tatagcaaat ltaaactatg cttttatitt 42720
 agagcatcac caaaagaita gcaataatit atccclaaaa ticaagittt gggttictta 42780
 5 aactgaaaai aggaagtga aaatctttc cgiccaagag atagcciaaa tcttatctta 42840
 actaatlaaa atattcataa ttttctttc gtcacattaa attttcgtcc glaaatccga 42900
 10 ttgaaatcca atiggacaat ccaaaaaata gagaaaaaga acagaaaaaa laataaaaag 42960
 cacacaaatc ttaictcaat cccgcgggaa gctgccgacg ccgccgaatc cgcicgagcg 43020
 ccgccgccgc cgcgcgcgt cacggggaac gatgtcgtc ctgtcgcacg cggtaiggga 43080
 15 gggcgccgc gccgtgctt gggagatagg atatggagag agaaggaaat gtagggagg 43140
 gttaggittt tccccatccg tatcttcagc gacacggagg cgaiccaagc tgcctatcag 43200
 atcggacggc tcagaacgcc tccatcgtca ggccgcgcat gcttgatggg ccgagggaag 43260
 20 gccggagggt cgaacaaacg cagtcagagg aggagttaga ggaggtaaag lagaatttat 43320
 ttgcgggctg agatagtaaa tggactgaaa atggcccata gagaaattgg gaattttatt 43380
 taaataaaig ttgaaaaggt gtttatatta tcaaaattag aaattaagct ccgaaaattt 43440
 25 taaaaaata tcaaagagca ttattaatca tgattaatit aataaaaaat aaatccaacc 43500
 atatcatait atttcacggc gcacggtagg aaaatgcga gctgttgtc ctgacggtagg 43560
 gagagaaggg acatigtita ttccagaac tatctttat aactcccatg gaactttaaa 43620
 30 ataaatataa tcatlattat agcatlagtt ttttctgtc tttttttcc ccaagagcgc 43680
 cgcgcagaag agatcatcg cgaicccct gccccgacgt cgcggccga tctctcattc 43740
 35 tctccacgcc ctgctcgtc ccgatctct acaccatccc tgccatctc tcttccctt 43800
 cccctctatc ctccactggt gccgccacc tctccgtata agacaaacig cgttgcggcg 43860
 ttggtttccg ccggcgctgc tgcctcacct gtcagctagg gcgggcatgg cgcgccgcgc 43920
 40 cgtttccgc gctgttggcg ccttcgtc ggacggctc atccaagggc gaggaggccg 43980
 cgcggggggc agtggcgccg aggacgcacg ccacgtgtc gacgaattgc tccgccgtgg 44040
 caggggcgcc tcatctacg gcttgaaccg cccctcgc gacgtcgcg gtgacagccc 44100
 45 cgcggccgcc gtgtccgct acaaccgcat ggcccgagcc ggcccgacg aggttaactc 44160
 cgacttgtc acctacggca ttctatcgg ttgtctgc cgcgcgggcc gcttgacct 44220
 50 cgtttcgc gccttgggca atgicattaa gaagggaatt agatggacg ccatcgctt 44280
 cactctctg ctcaagggcc tctgtccga caagaggacg agcgacgcaa tggacatagt 44340
 gctccgcaga atgaccgagc tggctgcat accaaatgtc ttctctaca atattctct 44400

caaggggcig tigtatgaga acagaagcca agaagctctc gagcigctgc acatgatggc 44460
 5 tgaatgatcga ggaggaggia gccacctga tigtgtgtcg tataccacig tcatcaatgg 44520
 ctctctcaaa gagggggatt cagacaaagc ttacagtaaa taccatgaaa tgcitggaccg 44580
 ggggattttaa ccigatgittg tgacctacaa ctctattatt gctgcgttat gcaaggctca 44640
 10 agctaaggac aaagccatgg aggtacttaa caccatgggtt aagaatggig tcatgcctga 44700
 ttgcatgaca tataatagta ttctgcatgg atatgctctc tcagggcagc cgaaagaggc 44760
 tattggattt ctcaaaaaga tgcgcagtga tgggtgtcga ccagatgttg ttacttatag 44820
 15 ctgtctcaig gattatcttt gcaagaacgg aagatgcatg gaagctagaa agattttcga 44880
 ttctatgacc aagaggggcc taaagcctga aattactacc tatggtaccc tgcctcaggg 44940
 gtaatgctacc aaaggagccc ttgttgagat gcatggctctc ttggatttga tggtagcaaa 45000
 20 cggtatccac cctgacattt atgttttcag cattcttaata tgtgcatacg cttaacaagg 45060
 gaaagtagat caggcaatgc ttgtgttcag caaatgagg cagcaaggat tgaatccgaa 45120
 tgcagtgcag tatggagcag ttataggcat actttgcaag tcaggcagag tagaagatgc 45180
 25 tatgctttat tttagcaga tgaatgaaga aggactaagc cctggcaaca ttgtttataa 45240
 ctccctaatt catggtttgt gcacctgtaa caaatgggag agggctgaag agttaattct 45300
 tgaaatgttg gatcgaggca tctgtctgaa cactattttc tttaattcaa taattgacag 45360
 tcattgcaaa gaaggagggt ttatagaatc tgaaaaactc tttagactga tggtagctat 45420
 tgggtgtgaag cccaalgtca ttacctacaa tactcttata aatggatatt gcttggcagg 45480
 35 taagatggat gaagcaatga agttacttct tggcatggct tcagttgggt tgaaacctaa 45540
 tactgttact tatagcactt tgattaatgg ctactgcaaa attagtagga tggaagacgc 45600
 gttagtctct tttaaggaga tggagagcag tgggtttagt cctgatatta ttacgtataa 45660
 40 cataattctg caaggtttat ttcaaaccag aagaactgct gctgcaaaag aactctatgt 45720
 taggattacc gaaagtgga cgcagattga acttagcaca tacaacataa tctttcatgg 45780
 actttgcaaa aacaaactca ctgatgatgc acttcagatg ttacagaacc tatgtttgat 45840
 45 ggatttgaag cttagaggct ggactttcaa cattatgatt gatgcattgc ttaaagttag 45900
 cagaaatgat gaagccaagg atttgtttgt tgccttctcg tctaacgggt tagtgccgaa 45960
 ttattggacg tacagggtga tggctgaaaa tattatagga caggggttgc tagaagaatt 46020
 50 ggatcaactc ttcttttcaa tggaggacaa tggctgtact gttgactctg gcatgctaaa 46080
 ttcatattgt agggaaactgt tgcagagagg tgagataacc agggctggca cttaaccttc 46140

catgattgat gagaagcact ttccctcga agcatccact gcttcccttgt ttatagatct 46200
 5 tttgtctggg ggaaaatata aagaatatta taggtttctc cctgaaaaat acaagtcctt 46260
 tatagaatct ttgagctgct gaagcatttt gcagcttga aattcttgtt tggaaattctt 46320
 ttctcttaca gtcctattag aggagggatc ttctctgtat gtgtaaatag cgaggtaigt 46380
 10 atgccaccctc tccgaattat ttttactgtg gtccctagac tgtaaacaag caattatgtt 46440
 atgcigtiga tgcagaaaa aacataaaaag ttgtctgta tctctactaa cggatcataa 46500
 agggaattgt gactggagtt tcaaacttaa tggctctagg cagtaatttt gacattagat 46560
 15 ccaaaacaat ttatagggtt tcaataaatt tcatctatgt gtactgttta gggttgaat 46620
 agtttgactt gttttttaa lgaacaaaag atatgtctga agctttgttc ttaccaaat 46680
 gcagtactga tcatcacaat atatttttta tggacaaga ttggattgta tagaatgggt 46740
 20 tctgactga ttatcttata tcaacgtatt attatgcaca tglactaatc algaaatata 46800
 tgaatggaatg atgtttctat ttacctgtgt gaggcagcaa ggagtggat ggataacacc 46860
 25 acatactccc tctgtcccag aatataagaa gttttagagt tggacacgat tattaagaaa 46920
 gtaggtagaa gtgagtgtg gagggttgtg attgcatgag tagtggaggt aggtgggaaa 46980
 agtgaatggt ggagggttgt gattggttgg gaagagaatg ttggtagaga agttgttata 47040
 30 ttttggggag tacattatta ttctagaaca atactgttgt gctcaagaag cgttccaaag 47100
 atgtttcaca accgtgtctc gatgggtttt gagcttaatc ctgggacatt cagtatcatg 47160
 atctgtctca ttcttaaaca tggaaataag gatgacagca tgatttcttt gtctctataa 47220
 35 tcttttggct acccacagat aatagctgta aatctatact acittaaaag gagtagtggt 47280
 ggtggtagt ggtgaatctg ccaccacccc accaccaact ctcaaaatc tgacatgttg 47340
 gatcactgtc aatcccttct ccaagacatg tgggactact gtcaatccct tctccaaacc 47400
 40 aattgtatga tagaacagtg gaaatcacgg acagaccatg gagctctcaa ccataatcat 47460
 ccttgcgagt taataacaaa tggagcgtaa acttggaag caaaaaactc aaattaatc 47520
 45 taaaattaag ctctaggatt caaatagat ttctctctg cattgtgtg ttatgatttt 47580
 taattccgta acaacgcaa tgcattttgc tagtcttata aagaagggtt aatgcaata 47640
 ttctgattaa atgattgtat ctatgaagtt tgaatgctag tggagctcc ttgaccaatg 47700
 50 ttttgttgt cgagcattta agagagtga gagaatgctt ctttgggtgt gtctgtgtat 47760
 ggaaggatcc acagataaaa ttcaggttct actgttctc tgcctglaat ttcatgaag 47820
 ctgcagtga taccttgttg accacttgat ctgttgcctt gaaggagaat atagtatgg 47880

ccaaggttgg tgacggtgat ggtggcatgt gatccccag atcttcagtg acccagagag 47940
 gaggggacgg cgcgttgtga gctacaaggc atactcagtg gagggcaaga tcaaggccctc 48000
 ccgtccglag gggactccgc tgcatacagg ccaactgctc cgaactgalc aatttcttgt 48060
 acggatcact tctcttttcc tttttttttt caccctaagc actctcttga ttcttcgctg 48120
 ctacctccct taatttcttt caatatattg tggcacttga tcatggcgga gaccacattt 48180
 ccagtgtaga tggattttgt caaagaacta aatttattcc attagcttat ttccgattt 48240
 catggaagac attcttttct ggaataaata cagaactaaa tctgttttcc tgaataaaag 48300
 ttgttagtgt gtggcatgtt gcatttccgc gcttctaat ttataaaaac ctgttcattc 48360
 aatttgaacc tgcattcaat ccaatatatt aggtgcagac aggtgcttgc ggtcaggtta 48420
 aagaagtigg caaaaatgct tctgaagaaa ggttaattgt tgtttcatct caggaggtaa 48480
 tatgcagatg attattccaa ttggcattgc ctggccattt ttatcacgag tctttacaat 48540
 ttatatacct cctacataat ctctccagat tccagatgat ccagtgctc caacaattga 48600
 ggcgcttatt ttgcctcata glaaagtaag tacacttgcg gagaaccacc agttgacaac 48660
 acggcttgtt giaccatcaa acaaagtigg ttgtattctt gggaagggtg gaaaggtaat 48720
 tactgaaatg agaagacgga ctggggctga aatccgagtc tactcaaaag cagataaacc 48780
 taagtacctg tcttttgatg aggagcttgt gcaggtaatt tatltggcca tacttacacc 48840
 agagatccat atattacttt tataactgca gtttttactt gttaacattt cattgtgctt 48900
 ttacatttgt tccaagcttt caggttgcgt ggcttccagc tatlgaaaga ggagccciga 48960
 cagagattgc ttcgaggctt tgaactagga cactcagaga tgggaagtct tccaataatc 49020
 cgacaccttt tgccttgtt gatggctctc ctgttgatat ctgacctaac aaggaaatca 49080
 tgctataatg acgacttgcg aatagtcctc calatggagg gcttgctaat gatccacat 49140
 atggaagacc tgcattgat ccaccataat gaagaccaat atccacaata tggagacct 49200
 gccaatgatc caccataatg aagacctgtc aatgatacat catattgagg gtigaacaat 49260
 gatgggcctc gtgactaggg ccggtcctga ggggggtcga atggggcgat cgctccgggc 49320
 ccccgattc ccagggcccc cacttactgt tgcaacgagt agtagcgatc ttccagcgcg 49380
 caacgtgagg cgaatttctt ccgtgatttc gccggcctgc aactgcgaga tgcgagiat 49440
 aacgatcagc cgatcgatct catctgccga ctgccatgt gatgccacac gcaagcgag 49500
 catatcagcc ttatcttgtt tgaicggcat gcggacgag cacatctgtt gtcgatcaa 49560
 ctgctgactg ctatatatgt gctgggtgtg aatcgatcga ttgtcgtcac ggaagtgaag 49620

aacaaccacg gcacigcigc cigcigggct ctagccgcca tcaglaagta cgctatacig 49680
cctatctaga tctagatcga gattacatag iggaattatc tgitlataac aaaattacaa 49740
ggtaicaatt gataatttaa gggtataacc glacaaactt cagtgatttg ctgggttcac 49800
attggllaga ttgttticaa ctaatttggg acttcigtag ccttglattt tacgaatcta 49860
gtattlaatat tttcttaagi attagcctgt tccctgatat tatgcigtig agaaagtag 49920
caatagataa caaaaacaag taggttggtt gaggatgctc aagagtaata caggcacttc 49980
aataattcig atattatcag gacatcatca ataattctgc gcctacaaat ctcaaaagaa 50040
aatitttaata taatgcgtat gattttttta atacgaatat tgattgctat ttaaagatat 50100
ttataattata tggtaattat tatttgaagg ttataataa aggcctccgt ttttagtttc 50160
acgctgggcc ttcagaatct caggaccggc cctgcctatg atccttacac cgtgtatcct 50220
glagagtact tctctaaaag agagtacctt agtggagta gcaaagtgc accatctgct 50280
tcatacgaaa gatatgcagc aactacigc ttgcctaata gagaacigcc ctcatctatt 50340
agtccctggg ccgattatat gtccctccgt tcttatcttg accaagtacc tactgatagg 50400
tactctaata gggttacact acaattaggc ctcttgagag ccgggaatag taatgtgcaa 50460
caattaggaa tcaccagagc tggaaattcc aatgcctatg attatactga ggtacatttc 50520
caatgcgtta gcttgccctt tctttgcaa tggccctgc ctgatagtt tccattagaa 50580
acatgaaacc atataattga ctgttgcat atgtctattt tcttccatga tggttcagac 50640
gtctgaaaaa aggacaaaaa tattctagaa tatgtcatgg tgatccaaat atatccttct 50700
gtcttltgcc cactctaata tctatcgttg gtaacactat tcaattgtta ccatgttgtt 50760
gcaaacccta gattcagttt ttacagctgt ctctgctgct gtgtcttacc agttttctta 50820
gttgggtgtt gatctttct catTTTTTat ttcttgttt ccgtgttcac ctgctgccctc 50880
tctgatgcat ctgaatgtat attttgttc tcttcagtgc ttaatagatt taaatttcat 50940
tcttttcagg ctgcggagct gatccatgga cgtgaggatt accgaagact gtcaggctctc 51000
actgggtatg gcttacgcag actgaatttt tacaggacac aaacatgaat ttgtcctca 51060
taatcattga gtgatgatct ctgtgcaggt atccagggtg ctctgtcgaa ttgttgattc 51120
caaatagtta actggagict gtcattgggt ttgggtgtgt caatctagct gagatccgtc 51180
tggtatagcg taagagaaac atcaltgact atccccagtc ataaccatgc cccaatggcc 51240
accaatagtt ttctcglga aaatctcccc ttgatccag atctctgtgt cgagagtgaa 51300
gttgcacgaa gcccatcctg gtcttccga gtccattgtg gagatccagg gcattccgga 51360

tcaagtgaag gccgcacaga gccctcigca aggccttcac ggcgcaagca gcaacagcag 51420
 gcaggcgccc cagtcctcic gcatggccca ttatttttag taagctggag gacattcgca 51480
 5 acaggggggt cagtggtcac tgcaaagcig agtttgttct tcagttcaac tgcagaaaaat 51540
 tgcagatcgg ttgccgtagi tgctagaacg gtacatagtt gccacctaac tglagcgagt 51600
 10 ggcataactt atttgttgtt actgcccaat gtgtctcic ctgtgttca tggattcaga 51660
 ctgtgtattg tagtatttct ggatcagact ggagtaaaag aaaaaaaaaa aggaagacat 51720
 gggtttaaca gtaagctcaa aacgttgaca gtagtaaaat aaaaggggtt tgttacttt 51780
 15 atttccaata tcaaccttac caacatttgg cgttgaatca ttataccac atcgcttgig 51840
 cagctgaatt tggggctgtt taaaagatgg tctcttggat tgctaatgc ctgcgga 51900
 gcgtgggtacc ttgtacaata taaatataat tataactatt taattcata attaaacatg 51960
 20 ttgttcaaaa tctctactat tataaaaaat gaagatgtt ttgtccggtt ttltgttacc 52020
 tcatctgigt atgaatccgt ttttaagttc gtttgccttt ggaaatcat atctgtatt 52080
 gattcagttt ataagatcgt tcacttttgg taatcacaga ggaatcata aagaattctg 52140
 25 tttaaaaaca ctgttatagi aacttgagac gatcagacgc ctaactacag ctcatgatt 52200
 tctaaatata tatatatata tatatatata tactagaaaa aatataatgt tgltaaaagc 52260
 tatcttaatc ttattattgt tatatatatt agttaacaag aaactatgt tgggaacttg 52320
 30 ttgttatata tattttttta aaaaaaatca tgagctgcaa ttaggaatcc aatcgtctca 52380
 agttagcagg agggcgagtt tttttaaaga gatttcttat acgatttctt ctatatttct 52440
 35 aaaagcaaac gaacttaaaa accgactcaa acatggatct gtatttccaa aaacgaataa 52500
 acttaaaaac cgactcatgc acagatgatt aatttttata atagtagaga taaacgaact 52560
 cccacagtga attttatttt aactgaacca tataacaata ataagattaa aatagacttc 52620
 40 accggttgca atgcacgggc attttttcta gttaaagaag aaataaaaaa acacaaaaat 52680
 ttataaaaat taaaaaagaa aaatattata attttgttag aattattatt ataatataga 52740
 45 aaaatagttg ccaaaatttc tcaacgaatg tcgaataaac tcagcaatgt cataatatt 52800
 aatatgatgg taatatttgt tcgcaaaact ttaatttca atcttcaac aacatagata 52860
 tacaacgtcg taatcgccaa caagcccgag tgaccataca ggatagccga gcggtggatc 52920
 50 tgtactgttc ttggglgaaa taaatctagt acattgtata tcttactta atactacta 52980
 ttataaaaaat tgaagatatt tcttcaaga ttccatagc ttcttactc cgttacaata 53040
 tcggttctac tccgttaca latcggtttt gtacaccccg cgcacgcgtt gtgtgttctc 53100

ccgttccaat acatgaagct agagicttgc ttctccctgg tctggcaggc cctttttcca 53160
 ccatccccac cagggccagc gggttacatt gaccgatcac ggcccacatt agtggatgca 53220
 gccagccacg cctttcacia atcatgtgat gaacattagc tgagttaaaa tttatccttt 53280
 gatgattgtt agaaatgttt ttcttccac atcttctctt tcaattttgg aaaaatagat 53340
 ttcttgattt ttgtgtctgt acalcactaa laaatcagtt gttacccttc cacacattgt 53400
 caatttacca tgtctatttc agctcttacc ttgtatagtc ttgactcttg agtccctcgt 53460
 attgactaag ttgctacatg cctcctacaa atcaatagac tgccataaca atattttcta 53520
 cgacatgac catattagtc catgcaatgc aagtacacac aactactgc acgaaaaaac 53580
 tatgcacat aacttcaaaa ctaacatgtt agaatgacgt taatttttca ttacaattat 53640
 attcatcgac cgtaatttta ctaggcatcc tgtttaaaaa aaatattcac cgaccatacc 53700
 cacatgttcc gtatgtcatt aggtagtga lcggtagtta cagcagctgg atttttatat 53760
 ttgtgtcatt ttgaaaaatt tatttcgcaa atagactcct gaaaaaactt atcccagaaa 53820
 tagtcccttt tggagcgta gagtggctgg cgccgtggc caacgggaca gcgccaacct 53880
 ctctggcgcc gcccccgcc tctattcttg ttctctata tagagtigca aactttttat 53940
 ttttgtttta tttttttgga tgtttttca ctcttagaat cacgatata ccaactata 54000
 aaaaaattaa actcgaacgg aatatacac ttagctagaa gtctgaaaat atagcatacc 54060
 acttatctac ttgcacctt caccaaaatt agaccataac ttctttagta aaatcctttg 54120
 atcagcatat taaacataat gcactctatc actaggtgaa attacttaat ctaattcaaa 54180
 atataactac atgtagcctt gaaaaattct acatgccaca tatttcgtcc gtttgagttt 54240
 attattttta tggttcgttc atgtgagtic ccaagtgiga aaaaaaata aaataaaaat 54300
 aaaaaagtig cacatccctt cctctgcatt agagaggaga ggagaggaaa aattctacag 54360
 gtcacatatt tcttccattt gatttcattt ttctatggg tggttcttgt gtgttcttaa 54420
 gcgtgaaaaa aatatcaaaa aaataaataa aaataaaaaa attcgggggg gggggcgcc 54480
 agccactctt aggggtgaaa acgatcgga aalatccgat ccaatctgct ccgaatccat 54540
 ccgaaataag gatatggat gggtttttag aaatctggcg gatattgat cggatagga 54600
 tatggatatc ccgaaatagc acggattatc cgacattttt gtcggattat ccgataggcc 54660
 ctltaccgga taatccgaaa ttatgaacac atgtaaccac tctatctatt gcatataaca 54720
 taagtggtc catccaatga cctaattcat caattaccct agatttctta ctatgtggtt 54780
 ttaccattt catgtcacac ttgcgtagct gtatttttat aaaatggaca tcatgtattt 54840

atgttgttta gcacctaagc acataattat tacaatgggt cgtttatiga catigtigta 54900
 tttttacttg catigttaac tcaatgttgt attgattgca tacacacgta acatctgata 54960
 aaatttaate cgtttctgaa cggattccgc accatttccg acatctgcat ccgtacacta 55020
 tccacaccca ctccgaatcc gcttaaaaaat atgggttagg ataaggtaag accactatcc 55080
 gtccgaatcc gctttatitit cacccttagc cactctggcg cgcttccctt gccacctcag 55140
 catcgtccca ccacgtcggc agaaggacgg cggctccagc cactctggcg ccacaaaaaa 55200
 ggaccatttc tagcataagt ttttttaggg gtctatttac gaaataagtt tttaaaagga 55260
 ccaaaatgtg aaaaatccag gtacacagc acigigataa gcaatagcta tatgcctat 55320
 atatacacgt atatgcatig ctaatccitc aattttgtcc aattctttta aattgtcttc 55380
 acctgttgca acgcatgatt tttttcttag tcttaacctt aactaatctt aataactaac 55440
 taaaagattc gtacttttcc gatcgtcacc ttgtccatac gctaattitit cgtccgtccc 55500
 ccttccctt caaaaaaaaa gggaaaaatc catittacac cctcgaacac ttatgcttgt 55560
 ctaaaataca ccccggaact ataaaaccgg gtataatata cctcagagct atcaataccg 55620
 gacagttcaa ggggtgtatta tacttggttt tgtagtgttg ggggtgtatt tagataagca 55680
 taagagttca agggcgtaaa ttgacttttc ccaaaaaaaaa atcccagtcg ttactttcca 55740
 tcttgagaat cggagacagg gaaaactgaa gcatacacgc aaatagaatc aaagataggg 55800
 aaaactaagc atatacacac aaatataatc aaaaattccc atgcagctag atcgggtgcc 55860
 accgttgttg ccaaacacc accattgcaat gtaaaatctaa gactaaagcc taaatcctat 55920
 gctaagtcac caaattagac tgggttctac caatttggtt atatatcaaa ttagacttga 55980
 tttttactga tttagaggttc tggagggttc acactatgaa acggaagttt tcccgttgc 56040
 aacgcacggg cactatgcaa tatcttaact aattaaaaga ttcatatitit tcttttcgtc 56100
 acaccgatct ttcttccgtc tgaacaatca cgtgcacctc ctctccaaat cccacatcat 56160
 cataatccga ccaaaaaaca aaatctcaat ctcaatccaa tcagaatcat caaaaaatca 56220
 tccaaaatat caagagatga ttataggaga tggaggggtg agcaggagca acatcatcat 56280
 cgcataaaaa ccccaaaatc aatcacaaca acgacatcat tatcacataa gaaaaacaat 56340
 acaaacaca tacacaatca acaacactgg cggatccagc cgaggggaca acggcgtggc 56400
 agcgggcaga tctctcgtt cagatccgcc caggggtgcc actgacgtcg ccgccgccac 56460
 cggatccaag ggagaagctt cggacagagg gagagggggg tagaggaccg ctaaatccgc 56520
 ccaccggaaa tgcgcgcc accacctccg tcggatttgc ccgagggagc gccgatgccg 56580

ccaccgccaat cgcgggagaa gcttgggcac ggagggtag gagggagggg ggtagagaat 56640
 5 cgccggatcc atccgctgga aaagccctcg ccggatccgc ctgccgaaa caccggtgic 56700
 gccgctccg ccggatcgg lagcgggagc cgccgatcc accaccgcc ccggatccgg 56760
 tgggtgggag ccactgacac catcgccgc gccctctctg ctaccgaaa gggagagacg 56820
 10 agaggggagg gggcgaggc gggggacgag agggtagag ggagggaccg agtgggagag 56880
 agagggacga gtgagaggag ggggacgagt gaataaggat gcgtgacctt atccactcgc 56940
 gcggtcgcac cccggctctt tctctcgctc agctgttcg ctgtggaga ggalgcgaga 57000
 15 tttttttt aglaaaatgc acgggaggc cttaaacttg tagcggtcgt tcatctaggt 57060
 tcccaaactc tcaaaatgca tatccaggc ctagaatttg tcaaagtgta tcatctagat 57120
 cccaaaccga cacatccctt ctgggatcc acatggcgct aatgtgactt gtcacatgga 57180
 20 cgtgacacgt cttttttt cttctttct tttctttc cgtttcttc tcatcttct 57240
 ttttttccat ctctgctcg ggacacatag aaaggaaaag aaaggaaaat acaagagaag 57300
 aaaaaagaa aaaagaaaat ttttaaatgg gtcctattcg tcagtcaaaa ttatgccaca 57360
 25 tcatgtccct gcgacatgcc acatcagcac cactagcat cctgaagggg ttgtggcgat 57420
 ttgggacctt aatgacacac tatgacaagt tctaggactt ggatatgtat tttagagatt 57480
 30 taaggattta tatgacacac tactataagt ttaaggaccg cccatgccct ttacttttt 57540
 tttttacacg gagagaatgc gaatttgttg gttagttagc gctgagggtt tctcgcacgg 57600
 agaaatttgc ggtgggagaa tttttttc aggttctttc tattgggaga agacgggatt 57660
 35 ataggattta ttactgggtt ggtggccctt gtttctttc tttttcgagc tttttccgt 57720
 taaattcact tttctctctt caaggagcgt aggacatgac tgaatgcagc tgctgtaaat 57780
 tagaaataaa aaagaaacat attctgtttt tcatlltttt caataggtaa atataaagat 57840
 40 ttttaagtaa tatttaaaaa tatatagtgc tgatcaacga cattgttaag tgagattttg 57900
 cgttactat cacttttttt tccattgggc tcacgtacgg cattaaaagt tttagttttg 57960
 45 gtctctctt tttagatttg ggcataacc aatattgaga taggtatct aaagttcatt 58020
 tggattttat tcgattcaac tttttgggt ttgtttcagt tctttttac atgtttctca 58080
 tctgaaatta ggaaattagg ttgggtaaag tcttgaatag ataacgctgt tgacgtttga 58140
 50 acatatattt atctatttat ttatttaaaa atatatgaat aatttttatt ttgttatgac 58200
 tttgtcggg gacatgggac cgggagttat atgactagag gcttgggcag gagcgatcac 58260
 ccacgtggcc tgatgtaaca tctgaaaa tcccaacaat aaaaatcact aaaattttga 58320

acttttttaa actttttgcat catgcctgggt gttaatgattg ctattgcttg ccaaaccgtt 58380
 5 aatgatacaca aagaaagtaa agtaaggatc taaaattttaa gtaatatagata aattttacgag 58440
 aatataaatat ttaattgcta accctacaaa taattacgca caagaaaaca aagccagaca 58500
 aacggaaggt taattactaa tttaaattat ggatttaatta ttaaatactt gaaccatgtg 58560
 10 ttgcctgcca tggcatctaa atacacatga aataatggc atataattaa attaagcttt 58620
 ataaaaattat gtgagggtttt aattaagcaa itagcttaat gtgtaccga gtccttaatat 58680
 actatittata gaataaataa attcaaccia tccgtgtaaa atatatgtct ataagttcat 58740
 15 tcaatgtact attgtaataa taatggccac attaggatat ttttaattaa tttggaacct 58800
 tcaaagcctc caaaattatc taggttaatt ttgaaattat acctattta agtaatgcaa 58860
 tagaaaaata tacataaaaa taaaatatgg gtaatatag aaattgagta aattttcatc 58920
 20 taaattaaaa catatatgg gtaaaccctc tttaatgtaaa aattaagatt tatagaatga 58980
 aatttgtaca agggataaac taaaatcggg ttaaatagaa aatggcactg ttcatgtcac 59040
 tctagggtgt ctagctggc cctggcccta ttttccccc cagccgctcg cgcttggtg 59100
 25 cctcgcgccc cgcgccacgc caccgcgctc gcgtcgccgc tgcgcgccc tgcgcgtcg 59160
 ccgttcgcgc ccgtcgtcc gtgcgtccgc cgcccgccgc cccgcgccc gtcgtcatcg 59220
 30 cgtcgcgctc gccatcacgc cgcttgccg cccctgacct cgcgcgccc cgcgcgtcc 59280
 cgtagccgc tgcgcgttc atcgccgtc cgcgcgccc cgccgtcac gcgcgcccgt 59340
 cgtccgccc gcatagcccc gcgcgccc gccatcgtgt cgcgcgccc tgcgtcgt 59400
 35 ctgagcccc gcatccctct cgagccccgc acgtcgcgtc ttgtcgccgt tgcgtcccgc 59460
 tgcgtcgc cgatgtgtc gcgtcgccc tgcgcgccc cggtcgcct tgcgccccgt 59520
 gccgcgctc ccgcgtgtc gtgtcacct tgcgtccc cctcgtgccc cgcgccaccg 59580
 40 ctgcccc gtcacgccc gtgcgtcg cgcgcgccc ccgtgccc gccgtcacgc 59640
 tctgtcgc gtggccctc cgcttgagc cgcgcgccc cgctccctc gcgcctgcgc 59700
 cccgcgcac ggccgtccc tgcgtcgc ccgtgccc tgcgcgccc cccgtccat 59760
 45 cgcgcgagc ccgtgccc cgcgcgctc gcgtgcccc gccgtcacgc ccgtcgcgc 59820
 cctcgagcca cagcgtcgc gccgtcgt cgcatlagg gccggccacc cttttcccc 59880
 50 cgccctataa aacccccgg ccaccccc ttcacccac accatccca ccatcccc 59940
 tcttctctc ctcttccc tcttctccc ctccaccgc cgcgcgccc gccctcgtc 60000
 cgccgcgccc tgcgcgtc tgcgcgccc ctgcgcgccc cgcaccgccc ccttcgtgc 60060

gccgcgccgt gcgccgacgt cgtgccgccg tcgccgtcgc cgicgicgic ccgccgtcgc 60120
 cgicgccgic gtcggtaagc cgccgtccct tccctcgltc cgacgccgtc gccgcccggg 60180
 5 tgggaaggag ccgagagaga gagggaggaa ggagccggga giaggaagaa agaaaagaaa 60240
 agagagagag agaaaagaaa agagaagaaa agagaaaaga gagaaaagaa aagaaaagag 60300
 10 attagagaag ggagggaaga gtgggccccca cctgtcatta gccccatcca attccccitta 60360
 gaaaaataat tcigttagaaa agaaaatcaa gatcttgacc ccacctgtca gtcactatag 60420
 cgtgtggata aggttgtatt aaaaataaat gaattaggaa cagtactatt tcgcaactat 60480
 15 tagaattaat tcaaatitga atctttacac tagcataact aattcatttt agctccgatt 60540
 tgagtggaaac ttgaacctaa attcatctaa attcalaagc ttccaatgg tataataatt 60600
 actattaaat aaaatataat tataattatt aagtaattaa tatcatatga ttaggttatg 60660
 20 gtcaacttaa aaataigcta ataaataaaa ttagtatgtt ggaigtataa atatttgtct 60720
 ctaacatgtc ttgccactgt aacaaccaca caaactaata ttaagtgatg tctgaaatga 60780
 atgaatgaat aggaaaatac tagtacttgt ttaatatcgc atagccaat aattaaaccc 60840
 25 atggcttata ggtaatttaa atcaaatgta gccttgtgat taigcaacta aaatataaac 60900
 acatatagat gaatctttag ctgtattagg aggaataata acagagctag tgtgactagt 60960
 tatgatatag ctgttgtcgc gttagctata ttagtlaaat ggttcaatgt taatacactg 61020
 atgcacacac ataccctttt tgataaccta ctagtgtcat atattaaact tggtaataaa 61080
 tgaagaacca atataatagc taaatactgg tgctagttaa aaatcttgac cacacataat 61140
 35 ttagtticaa accacacctg aggatgttc gtataaagt talaaagtta taaagttata 61200
 caaaagataa talgiaacta taataglatt aaaccacaaa tctaaaatac agggcgcata 61260
 atgttcaacc tttatgcaa acggataata tccatatata tacatcatgt ggataattcg 61320
 40 aataatagct ccatggtaa aataataatg taggcgaatc atggtgatga gatggtttat 61380
 cctaaacctc cccatcgaca tagccatgtc atagggacct gaccatttta ccttcataac 61440
 agatctcttc cataagccaa tagctagact aaaccacaga ttagcaaatg tgtacatcat 61500
 45 atatttgtct agttagtacc aatagaacca tcaggacaat alaaatacta aggaatctta 61560
 gctcttagct tgattagaat ccaatagcaa acacgaglag talgagcagc cttaggttcg 61620
 accicaataa ttatattitg ctgttgcata atgtctctt gtigaatatt ggttttctc 61680
 50 gcatattata gaaattgtat atcggttagt cgtgaggcaa cgtatgcagc tttcaggagg 61740
 tgaaggtilga tcaagattgt atcaagaata atgactattc taagcaggca agtcatcact 61800

attccitgaa catgttgatc ctaatigcga aallatitig tttacaaata aaattgcaig 61860
 caatgatgaa catcctacti gtgattatgc catgccitga ttattgitta cccittaaaat 61920
 cctitgiaacc atgattacgi atgagiccct agtcaattat gacaattgct tagagatgct 61980
 attctagaat catgcataci catattitac aaatgctata tgcctgggca attaccttgg 62040
 ggaaggtaat tgagatgcgg catgtggaga calgaacgcc acattgccat gatattaatg 62100
 acaigattig tgaaggaga aataaaatia aacaactgii ttcgaciggg gcggacggag 62160
 gatttgggtg giatciggaa aaggctagta ccgtccccgg tcaaltaagg accgagccat 62220
 gaagttaagc atgaaacgac ccccgtaaca ccgcactici cgtatgggia tagacctagc 62280
 ggagtagata gctgagcgga ggcagiatcc atgcatagtg gtctcttgat gtgtgaggca 62340
 ggggctctac ggtggggcag ccatlggtag gaccgcaagg cggglatcta caglggtgic 62400
 gccatcggia ggacigccat gtgagaatci aaaacataat talaacttaa tgcattgltg 62460
 agtcttccct tcccggtgic gccagaactc ctctcactgc tagaaaccgt gtacgcctag 62520
 atgcatagag gatgaaaagt tcatggagcg ggtactgcca atgcgaggii atcgaaaagc 62580
 tctgccgtga cgcattctat gtgtlgggac gaggctcatg tgttgggcag tcgcggaglg 62640
 cgggtaaatg gtacatccac tgcagtiga gtaaaccaaa tctattcgaa tagccgtgct 62700
 cgcggtiatt gagcaccggg acatgtatta cactlggcta gactctaat tcttaacttg 62760
 tggggaatgg gatattgcat gatgaatitii atgctgatgg agccacatcc cgagaggagg 62820
 gaaggltggac atcttcagaa aacatgacg attcaatggc gggaagctat cctlgggac 62880
 acaatggatg gtggacagaa ccgtcgttgt ttaaagttaa cactggctact aaaatttgat 62940
 cgatctatgc taggtititg gctltgtaaa agaattgtaa aattagctit atgcaaaagg 63000
 acctgaagcc attccitgaa atacctcta tcatatgcat tgttattatg gtggcttgc 63060
 gagtacggtt ggtactcacc ctgtctatit atatactit taggagagt tgaagagaa 63120
 gcccttgcg gtacgttgc gtatccaca agatgatcgg agtgcggtct tgttctaggt 63180
 ctctgttccc cagtcgactg cctltggcat gtaaccggg ccttatatit atltgtctit 63240
 tcgtgttgt tctctgatag ttgttggcct acctggccct aatgtaagta ttaactctit 63300
 ttagcctaaa ttcattcgtg atatgttgt atccaactat gtatgttgt accaactact 63360
 gatccaggga ttggtagga taaacacaga agatttccga ttccaaaat cgggggicta 63420
 cactgaccc cctcaggggg ggggggtcgg gcccgagggt gatgtggccg cccctctit 63480
 tgtctccccg aggggtcgg cgcctccgt tctgccccg agggctgagg cgccccgacc 63540

ccttgggggt ttggcgccgc gtgtatgggt taggtgagca caacggggct cacctaaccg 63600
 tatttatgtt ggtttggacg agcgcgicac gccgcatgta gcgcagtgca gcgcgctcgt 63660
 ttaiccggtc tgtgaccagt cacagaccgg tcagatcgtg ggtaggtgg caacaggcgg 63720
 tctgacacac gccicgcccc atcccgicag gataagagcc tccaggcact tgcacctagc 63780
 ccggagccag catgctaact cctggagatg acacgttggg cccggtcaga tataigccag 63840
 gcttcatccc aaccattaca agcaagatai tglatgaaga agggcgaaca tgcagattgc 63900
 tggactgaca cgtggiggac aagaatgacc gatttgtgac cggcttgaca ctggctcatgt 63960
 cgtcggcaga caaccalgtt cccacgttgc acctgccttc ggccggagtgg aggtagglat 64020
 gggccatccc atcagaaggt cgttcggaca gcagccattg caagtctccg cccatttatg 64080
 aagagatgac agggatgacc cctggagaga aaaaaaggag gaccttgccc acctaggagg 64140
 tgaggacgac tggaagggga gaggatctgg agagtatgac ccacgagagg aaaaaaggga 64200
 gaagaggggt tctagagtaa gagctctctg actctccagc tctttgtagc ttcttcgtac 64260
 acagatccac cagaaaaatg gaggtaggta ttacgtctct cagcggcccg aacctgtata 64320
 catcgcccggt gctttgtgtt tttttcattc tcgcgaactt tccacagact aggagcttag 64380
 aatctcgccc agggcccccg gccgaaccgg caaagggggg cctgcgcggt ctcccgtga 64440
 ggagccccac gctccgtcaa ctttggctta taattaaaaa tactctaaagg atattttttt 64500
 atattttatt ttcttatgtc tataatgaaat tttaaataag atagaatgggt aaacataat 64560
 tggaaaaaca tataatccaa agtccactat cacaagcgta gcatagatag gattacaata 64620
 cgtttccgcg aagacigtgt atacctactc tatccctgtt tccctgtgcg gttgtgcca 64680
 ttggggctgt tttttcatct cggattaact cgcgtggaaa ccgcgagacg aatgttttga 64740
 gccataattaa tccgtcatta gcataatagg gttattatag cacttatggc taatcatggc 64800
 ctaattagac ttaaaagatt cgtctcatga ttacatgca aactatgcaa ttagtttttc 64860
 tttttatcta tatttaatgc ttcatatatg tgcctaaaga ttgatgca tgttctggga 64920
 aaatcttttt ttaactaaac atgccaagg tgtttctcca attaatgga cccaaatca 64980
 ttccggctca cctttgtctt tcaatttct tccactaaa ggtgatgaca ctgacaaaag 65040
 gtccaaaagc tacaggatct gattttgtt catccatctg tgatgtgtcg gcaagccatc 65100
 catggagttc atccactcaa ctccctctc tcagagagag agagagagag agagacagac 65160
 agacacatgc atgatagatt gtgctaglac ggtagtaaca ttttatggc tcttttcta 65220
 aaattctagg ttgtttggaa aacaaaaatt ctgatgtgt caataaatta ataattatag 65280

gtatttattt taagtcactt taggtgttaa ttttgaatt ttaaactgct taaactctct 65340
 5 ttcgacgcat ctgagagcag gtacaatagc agactataag ccagctataa atatatatta 65400
 agtagataaa agaggaaaaa taagagtagc gggctataga ttigttagaca gctgcagcgc 65460
 gagctccaag atacataigt gtatgacatg tgagaccaa cattaattat gtaglatatg 65520
 10 tttataigta tctatigtat gaattggcta ttaaattgac talgggtgtg ttcggagggtg 65580
 ggtgttggga accatctccc aagcacggaa aacggagcgg tccattatgg cgtgattaat 65640
 taaglattag ctatttttta aaaaaataaa tcaatatgat ttttttaaac aacttttgta 65700
 15 tagaaacttt ttgcaaaaac tcaccgttta gtaglttgaa aagcgtgcgc gcggaalatg 65760
 agggagaggg gtgggaacc tcccatccg aacgcagcct atacatgatt tggagccaat 65820
 agttggctat aatattaaac ttgctctgag tggctcttga atcatcgaag tgatagaaat 65880
 20 catatgcaga aatgtttata ttigtgatgt aaaatttgaa tctaaaatta tttatatattt 65940
 gaaatggagg aagtactacc taaaacaagt atgagaaaga gacatgaaaa acacaaaaatc 66000
 tagacttaaa aataattgga attactagca ggaggctgaa gtcaatcaag acggcgaaga 66060
 25 aaagcacagg ggacagcaga cacgttaaca cgtaagtaaa caaacaagtg gttaatlaat 66120
 tagggggccc tcaagtcicc cctaaagcca ctaaacaiga caggtttggtg taccatggaa 66180
 30 aaaagggtga agcaaaactt tattctctct ctcatlagat taccagttgg aaagcaatcc 66240
 tgggacctct agctaactct attattgtag aacaacgttt tcttagagag agagagagag 66300
 agaaaataagt caataaaaaa tactactaat ccacttgaac cagtctgtc ggtgtcggat 66360
 35 gatttaccac attigacgaa acggactatt taticgacgt ttcgaaaaac acactttttt 66420
 agaaaaaaaa aacttccctc tattagccac tcgttttagt tatataccta tccgagtatc 66480
 tgttaagttt atttatcaaa atatttaatt tatctctata attaaatata caatccgtaa 66540
 40 aaacaatcac gcagtaattc gtttcaaact gagcctcagc tagaaaatca aaatggaaat 66600
 gaataacaat agcaacagta gagttagttt ttcggcttat catccgcaac ccaaattgca 66660
 attttaaact tagccctaga gttaattttt aaggcttggt taccatactt cattttccca 66720
 45 gcattagttt cttttgtcac taaaaatgt tttttaagt tgtttcgtc attttctcac 66780
 ggtttatcag cagtagagcg aagccattct tggagccigt ttggcacagc tctagctcca 66840
 50 gctctagctc cactcttctt ggagctggag ctgagcccaa cagttttagg tgcacaaaaa 66900
 ttaggagtgt agttgggtgg aactctctca caaaaaatgt tggagctgga tttagacagc 66960
 tccacaactt cactccaaac ccaactcctg aagttaaat gataagtga agctctatct 67020

atcaagccct ttttcttgat catgccttcta cctactccat ttttgcttct tggccctcac 67080
 aggaattgga aaggaaaggc gtataigcat caatgcatgc atgcgcacat caacctcgtc 67140
 catcaacctat cataatcatt atcatctcgc cagctgacga aaatgacctg calccatcca 67200
 tcacggacaa tccaagcgaa caccgctacc aacatcacag ccaacctgtt taltactagc 67260
 tcttgatacc actcctacat aaacactacg cgcaggctaa ttaattaagc gtgattactg 67320
 aagtaacatc taatcacgtc ctggcttagc ttaataaga caacagttag agcaggctaca 67380
 atagcagcag gatataagcc agctataaaa aaagagagaa aagagcaacg ggctacagat 67440
 ctatagccag ctgtagcatg gacttcaaga cacaacgtgt gtataacagg tgggaccaga 67500
 taataatagt gtagtatagt aagtaactat tatatatatt gactatagat gatttggagc 67560
 tattagtgtg ctatagtatt aaacttgcct atagagcagg tacaatagta ggataattagc 67620
 cagctataaa catattataa tgagataaac attgatagag aagagcagcg ggctacagat 67680
 ctgtagccag ctacaacacg gactccaaga cacaacgagt gtatgacaga tgggaccaga 67740
 tattagtagt atagtaagca actattatat aaattaacta ttacattggc tatagatgat 67800
 ttggagttag tagtgggcta tactattaaa ctttttctct tagcaaaaat caagcgccta 67860
 atcacattag aggagtagct ttgagacaaa ccaattagcg gcgaatcaag cgatctgcgt 67920
 ggctctacag tgalggggcg ggccggggcc acagcccgac agtgacaggg ggcttgacgc 67980
 atgtcagccct cagccctgga cgggagctag ccgttgtgtc cccgggggag gggagggggg 68040
 cattcccatc atttcgcccc tcttccgggc ccacatctca gtggggggtaa aggtgtaaat 68100
 tacttgcgacc gcgagctccag cgagcclaga ttggacctt gtgtccgttt gactgaaccg 68160
 gagctactcc ccaatacggg gggattgcgt tgtgtgcatg ccatgtgggc ccgagcgccc 68220
 ttgttctgtg gctttgggtt ggaaagggtg ccgtgtgagc tgtgcggtgt tgtactacgt 68280
 attagtataa atcatttttg ggtactactc cctccgtcca aagcttattt ataatttgtt 68340
 gtactccaac cgtccgtctt atttaaaaaa aatataaaaa aaattaaaaa aataagtcac 68400
 acataaaaata ttaatcatgt ttatcatct aacaataaaa aatactaatt ataaaaaat 68460
 ttcatataaa acggacagtc aaacattgtc acgaaaatct aatgtttgcc tttttttta 68520
 agaccaaggg agtatctacg aacaaagata atacatgta taatcatgaa gcccatgatg 68580
 tgattagccc ggccgtttga ctacccac gagctacgtg gctgacaagt ttaacttgtt 68640
 aactccatca ttctggatac tttagagcatg tacaatagca gactattagc cagctataaa 68700
 catattttta tgggataaaa gatgagagag aagagcagcg ggctacagat ttatagccag 68760

ctgcagcacg gactccaaga cgcaataigt gtaigacagg taagaccata tgtaataagt 68820
 atagtaagca actatitiat aaaciggcta ttagatcggc tatagataaa ttggagctag 68880
 5 tagtggacta tactattcaa ctigcictia tatgatataa atatigatat aactatatga 68940
 ttttgtaaat gacatgiltg tttatggatg gaciatgtgg ggtcggtcgc ctccgtagct 69000
 10 gaccaaaaata caaacitaaa acccctatct aaaaaaatct aactittgtt tataaatata 69060
 gatataaaaag ttcataatta gagccitcic ttttaaacga aaagagtact atgaaaacaa 69120
 ctcgtaatac aaagactaat tacgacgaaa agaaaatagt actgacaaga ggaaagcagt 69180
 15 gaacttgcct actccctccg taaaaaaaac caacctagac acggatataa cactatatat 69240
 ctagattcgt tcgttgtaat gaagtgicac ctccgtatct aggttgggtt ttctgtacga 69300
 aagaagtatg agtaaatcta aagctatgta tacccttcgt caaaaaaaaaa aagtaaacct 69360
 20 tgtactggtg cgtgtcacat cctaataataa tatgtttttt tatggagggt gtacagttga 69420
 aaaaaattga tgtgttttaa ggatgaaaaa tattggtlaa gtltggctatg taactctaga 69480
 aaaaaaaatg cagtaataat aaaatgctaa ttigctggag tactagatta tagacaatcc 69540
 25 agtccaggac acgacacctt cctacitctc tccacttcca ctctcaccgg ccaccgcgcg 69600
 ctctctctct ctctctcccc ctctctccgc aagattcttc ccccaaatcc caccgatcc 69660
 accgccgccc cccgctcgcc ggagtcctat cgtgccacc gccgccggag ccgcccgcgc 69720
 30 acgcccgcgc ggctgtctg ctgtgtgtgt gaggaggtag agttgtctgc gctcttccc 69780
 gcggccacct ccgctgtctg ctgtctctgc ttccgttggc attgcgggga ggtctgtgtc 69840
 cgggggacgt gggggctcgt gttggagcgc ggctgccgtt gaggtagggg gtgcggcgcg 69900
 gcgcggctcg cgtctgtctg ccggtggcgc gggcgcgggg ggaagcgtac gggggagggg 69960
 gagtgtggcg gcggcgcgcg gcgggtagg gacggcgccc gccaccacca ccggtctgtt 70020
 40 cgctggcagg cgctacgctt ccagatccgt acgccgttat gcttctctc gccgaactc 70080
 tctccatttg attagtaacc cctcgccgaa acgaggcctg tgaggcgccc gctttctggc 70140
 tggcttccct gtactcgctg ctgtctctg cctgttgggt taacctgtt ccatcgaatt 70200
 45 tgggtlaagcg aaacatcgcc tcatalgggc atttggggtt ctggcagcct taggtctgcc 70260
 atccgtcgcc gagcttccaa gtgaccggcg ctgttggta tatttgcttg ctgttcttg 70320
 tttggtaggt gcgctaaaac ttltgtctg catgaattt atgccacca talacagcaa 70380
 50 attactgagc tgaataaatt cggctaalta ggtccagcaa tatgacatct cgtggattga 70440
 atgctaagct gacatgttat cactgatgct ggcttatata taggttgttg agaagtgaag 70500

atgtcgacag gigaacccct gcgtgcagag ctatcatcca ggacgccgcc tticggtttg 70560
 5 aggctatgga ttgtgatgg aatcaglati lgggtggiga tcttctttat actaggtttc 70620
 atgtgccict ggiccatata ccgaaggaag ccgaagaagt ccttggataa gattccaglia 70680
 tctcaaatcc cggatgttcc caaggagatt gcagtagatg aagtictga gcatgctgtt 70740
 10 gtcgaaaact tccgtgtgca agaaagccac gcgatatcgg lgcaggagaa acattacgag 70800
 aaagattcag ggaaaatgct ggcacacttg gttaggagta aatcgagtga tgcgataat 70860
 ttgagccaat gcagctcggg gtaccaatgt gatagggtcg gtagctcgta ttctgggtgat 70920
 15 gaaggcagct cgggcaatgc taggaggcac ttttctcaat atgcaactgt ctcagcatcc 70980
 cctctgggtg gtctcccaga attctctcat ctgggtcggg gtcaattgggt tactctgaga 71040
 gatttggagc atgcaacaaa tccgttttcc aaggagaatg tcattggaga gggtagatat 71100
 20 ggggtagttt accgtggctg acicataaat ggaacigacg tgcgaataaa gaagcttctt 71160
 aataatatgt aagagatcct gaaatctatt ctgcgtttta cagaacttgt gactccttct 71220
 gatgccatca talttaattt ctttggatat gggtctgcag gggccaggca gaaaaggagt 71280
 25 tcagggttga agttgaggct attggccacg tcaggcataa gaatcttgct cgccttctag 71340
 gatatttgtt tgagggaatc cacaggtaaa gctatttatt aatcaccttt gctgatggat 71400
 30 ggctagcttt tgtttctact ggcacattat ttacttgcat agggatgtag gattgctctt 71460
 ggctatgtc cactactca ccagattatc tcaagggata ggttattcct gactgcactc 71520
 ctatgtctat cgatttttcc ccttccaaat ctgatggigg gattcagcat gcccagtgac 71580
 35 agattatgct cagtcacag aaaccttctt tggaccacca ttcttttacc atgaaaatgt 71640
 ggccatagct ccgaaagcta ggattcacta gaagcgcaca actgcctatt ggtttgtag 71700
 ttggctataa caaggcttta ctgaaatgta ctccatagt tcattacttt gtgaatgcct 71760
 40 gtcttgttc ttacgttcc ttctcatgca tgttcaaltc taaatttga ttcatgatat 71820
 gtccaagcta ctgtattctc caaagaaaaa cagaagtcca ttacattatg tatttccag 71880
 ttttccgcca ttttggalac tgcctagaa acaagttaat aatatagata tttatatggt 71940
 45 ttggccagtg ctgcttaagt gaccatcgag atagaaatig cttaagaaat atactaagat 72000
 gttgagtgtc aggtgttctc ggataatctt gttaccaaca aataggctct atgaatataa 72060
 50 tgggtgtctc ttacgtaat tcaaaatcca cactcagcca aaataatctg caatagggtg 72120
 ttgaaaatat gattatgttt ctcccttgtt ttcatcatga ctacagaaat gaacaatgtt 72180
 gctacatctt gtaataattt gtggtttcca attgaacaaa acatccatca atgatatct 72240

acagcaatat atttgcact tctgagcaca caatagggtt gagtglattc gagtcatggt 72300
 5 cattgattta agcttttlat ttactacat aaccattgat ttgagtgtat ctaaggagtt 72360
 ctgtttccac aagtacttta tgttaatggt gtctccttat gctttggcca tccaaactca 72420
 ttactgttgt ttaatatatt tagtggttag tgggtgccaa atctttcttt gtgtacatca 72480
 10 tactatgttt ttgtagicta ttaaacttcc atcctatcat ctgacttgtt atattccagg 72540
 atgcttgtat acgaatatgt gaataacggg aacttagaac agtggcttca lgggtgccatg 72600
 cgccaacatg gtgttcttcc ctgggaagcc cgaatgaaag ttgttcttgg aattgctaaa 72660
 15 gcgtaagaaa caaaccatcg tccccgtcaa aaagaaaaga attgttcttc actttagctc 72720
 ttttataatgt atatgtttag ttgcataacc cattttccat aactgaattg gtatacaggc 72780
 ttgcttattt acatgaagca atagagccaa aagtgttaca ccgggalatc aaatcaagca 72840
 20 acatactaata cgatgaagaa ttcaatggca aactttctga ttttggcttg gctaagatgc 72900
 tgggtgcagg gaagagccat atcacaactc gagttatggg aacttttggg taigtgtata 72960
 ttttttggga gttagtatta atctttccta tgcctagctt ttactgttgg aatgtgcagt 73020
 25 acttcgctta ttcatacagt ataaaatttt acatgctgcg aactttgtcc ttcgatatatt 73080
 ataacaggta gccttctcat tgctatcatt gattcatttc aggtatgttg cccctagata 73140
 30 tgccaacaca ggctctgtaa acgagaagag tgaigtctac agttttgggtg tgctattact 73200
 ggaagcagtg actggttagag atccagttga ttatggccgg cctgctaatt aggttagcat 73260
 atatcctaca atctcatgcg tattaigtat gttacaaaag tccgtactat lggaaattat 73320
 35 ttacggcaa aataacgtct atactaggag agacgaattt gcttcagggtg tatggctgtc 73380
 tggcagttgt ctactgtcta gttacccctg tctcactttt acagtctatt gttttatttt 73440
 tcaggagctg actagctgta taccttgtca tataatacaa cactgtaacg tggatgcctt 73500
 40 gcagggtcat ctagtggagt ggctcaaaat gatggttggc acaagaagag ctgaagaggt 73560
 agttgacctt gacatggagg tcaaaccgac catlcgggct ctaagcgtg ctctcctagt 73620
 45 ggcactgagg tgcgtcgacc cagacictga gaaaagacct actatgggtc atgttgttcg 73680
 gatgctcgag gcagaagatg tcccatcccc tgaggltgta acgttttctc ctltcctgca 73740
 ataacattca tcataattata tcatgtcaat aaatctgaag cttttgtctg aatcctactg 73800
 50 aaggaccgga ggagccggag gggcaacact gccaalgcag ataccgagtc caagacaagc 73860
 tcaagcgaal tggagataag tggcgataga agggactcag ggccatcagc aaggtttcaa 73920
 ctctaaagaag acggtgatca tagtcaagaa caatggcttc aaaactctat gcagtaacat 73980

5 ggtaggttggc agagaaaaag ggglatttct ggagggcatt gcaitttgta ttgtaggtct 74040
 gcatggcggg agagactgga gagagcacag tgctcgaiga tggatacccg gagacctgta 74100
 attcccattc agtattctgt ttgttagica agcagcttgt acagatcggt gtcgttcca 74160
 ttttttcatt ctcttggttt ttttgittag gaggcctctg gattaccagt acgaaccgct 74220
 10 gtctcttttc tagaatcacc aacatggaac ctatcaatat ttactactag tactacgact 74280
 tgcittcttc ttgcigagal ctatcagta ctgtacataa ctgacgtgtt cagctgcact 74340
 tggacaagta gatgctcggt ctgtatgtcg aatttacttg atgaggtcga gcattlaagia 74400
 15 ccatggctgc agccggcttc tgtttagtgt tgctgacatg cggcggcgac ctacgcctgt 74460
 gtggcccatc ctgatcttg ggccgaaact gtagcaacgg gcgtacggcc catclataac 74520
 gggattgttc ggcccgttgt agatgggccc gatcgggatt gcgacttacg tgcgacctat 74580
 20 ttccggttggg ccggtggctc gctacttcat ctagcagttg tcggcggcag gggtcacaat 74640
 tccaatagaa tccaaacatt attggattga gtlaaaaaca caaaccaatc ggctttttgt 74700
 caggttcaga aaattttaaa ctgaatttta attttttgac aaaaatctat ttgatttcg 74760
 25 tctgtttttt taggtttgtc aacggattca gcgaaatccg atgatatcgc tctgtagttg 74820
 atttttgatc cggatcagag attgtgaacc ctgtcgcgc attgcctgac aaagacaacc 74880
 30 agtgaagcgc cgtgcgcgcc gcgtgcgcgc cgcgtgacgc gaagatgcgc aggaaggaac 74940
 aagctggcaa gcggcgccgc catgacggcg gcggcgacga cgaccgcgc gcgtgcgtgc 75000
 gtcaacgcac gcgaccggcc gagatccgtc atggccgcgc gctatatata atacatcgtc 75060
 35 gccacacacc cccacacac cgagtcacgc ctgcgggag ttgagtttcg tagcggcgaa 75120
 ggatatagcc atatatata gatggcgatt gggttgggtg gctgtctgcgc cgtgtgtctc 75180
 gcggcgggcg tgccttctc ctctccggcc accacatgta agcagccca tcttctctt 75240
 40 ctctctctt tttcttctt ttttttttt tttttggaaa tgagccgcag ctgacaaaaa 75300
 gatcactcac acatggatac actgtcgtga cactaacca tgcctaagcc attttgttt 75360
 ctgtttttgg attttcttt ttatgtgtat cacttttgc tgtgtctctt gcagatgctt 75420
 45 atgattccct ggatccaaac ggcaacatca cgataaaatg ggatgtgat caatggactc 75480
 ctgatggcta tgcgtgaagt agcggltgca gtacaccaac atctctacct ttattttctt 75540
 50 ctcaacctgt acatttacac tatcttgtc tactacctct aataaaaaaa tatatttgat 75600
 gtlttaaaat ctattlaagt ctagagattt ggaaagctac acatggtttt atgtttgat 75660
 actattaagt agtatatttt alaagttata ttgaaggctg gggtttcaaa agtttgacta 75720

cactagatct tattcaaagc gtctaatgat tactgaacgg aggaagtatg aacttataga 75780
 ctitgaagtta aacagcatag ccacatctct tcatgtatac ttcattccgtt tcatattata 75840
 5 agatttticta gcattatcca tattcatata tgtgcgtcta gattcattaa taictataig 75900
 aattigggcaa tgcataaaaa tcttataacc tgagaaacgg agggagtatg tgcgaaacaa 75960
 10 caacaacaat aacaacgagc aaaatctgia tcgaatccgg ttccctctt gtaactgiat 76020
 caaagatcig tccctgtaa cgccccigt tcatcaggcc gttgtcacac tgtccaacta 76080
 ccagcaattc cggcacatcc agccaccggg gtggcagctg gggtaggacat ggcagcagaa 76140
 15 ggaggtagatc tgggtccatgt acggcgcgca ggccatcgag cagggcgact gctccatgic 76200
 caaggagggc agcaatgtcc cccacagcig caagaagcat cccaccgtcg tgcacctct 76260
 cccgggcacc ccaatcgacc tgcagatcgc caactgcigc aaggctggat cactgagcgc 76320
 20 attcagccag gacccggcaa attcigccgc gtcgtttcag atc 76363

25 <210> 28

<211> 53905

<212> DNA

30 <213> Orza sativa Asominori

<400> 28

gatcagtgag tgagagigat gigtatiga ttitcgtcta ggaatttgct gtgtcttct 60
 35 tcttcttctc cttctiacca agaaagatcg atggaggaga attttagga cgcgtttctc 120
 acgaattact tagctgttaa tgatcagctt gatgtgtacg atatgatggt gcagagtga 180
 agttgtgttg ttacatggig gatcatggga tgggaatatg ggaatgtgt aagaatgaac 240
 40 tcaagtgttt tcttttttgg gattactttt ggtaataaga gcttgggtga tgcgaaacta 300
 cagalggttt ttcttttaag ttgtatgac tctgtagagt ttitgagtaa ttgtagtgt 360
 tgtacctat caaagatcat ctttagctgc cttgagctc tccaactcta tatgtccatc 420
 45 tctagtatat atgtccata ttctgactg aaaattttca agtcggttgg ttccctccgc 480
 ctggatatc tttagctaa tttagatttt tttaaattgat aaatttgcta aaagcttgtt 540
 50 caaattcagc taagatctat tcaaatcca atttctctat cgaaattccc ggaaatttca 600
 attcaatcat tcccaatac atgccgattt ccgtaatat gaaccatgac atgtaaacaa 660
 cgaaggaac aagggcata ttagtttcat ctacatcga atatacggac acacattiga 720

agtattaaat gcactctaai aacaaaacaa attacagatt ccgccagaaa actacgagac 780
 5 gaatctatla agcctaatta atacalcatt agcaaatgtt tactatagca ccacattgtc 840
 aactcatgac gcaattlaggc tlaaaagatt cgtctcgag ttctctgacg aaccgtgtaa 900
 ttattatitt ttctacgttt aatactttat gtaatgtgcc aaatattcaa tgtgacaacg 960
 10 tgaaaatitt tatttggaaac laaataggcc ctaatatctt ttcaagatal tagaatagtt 1020
 atccctctcc accctccctgc acaaacagtg aacttctttc tcttgggca caggagtagt 1080
 agcagctccc ggaaacagaa agcaatcaag caaagtcctg aacctgaagc atcctgaaac 1140
 15 cagcagacgg cagaaaccag tgggcgcagg cgatagcagt ttttcgttgt ccggcgtaca 1200
 gccaaaat ac tggccatcgg gtgcttacct agaattgagtc cactggacgc agctaccacc 1260
 gtgtgtgcta cactgaccgc cgctgctcgt cgaccagttg tacggggctg acttattctg 1320
 20 aatttctaai ggittatttg ggggtttaga acactgaggg gtgctttaga tccaaagatg 1380
 tgaagtttgg gcgltgcaca tgggttatta tatataglt cgcacagggt gtttgggcac 1440
 taataaaaaat actaattatt gatcctat ac gataagctat ataatactcg atgtgacacg 1500
 25 ccaaaacttt acatccctga atctaacac ctttttaaat agagtatttg gtgtgaaata 1560
 taattttgat ttgggaagaa ggtgagtgag atttggaaaa aaaaagcatt tcaattaaaa 1620
 30 aatttggcag cagtaaataa agaaactact cggttttgta attaaagtga ggttttggca 1680
 ctcttttggc ctaaactggc ctccatttta taaagtgaga accgtgcagc aaaagcctga 1740
 aaaggcaaaa agaaagaaat tgltagaggtt tticaggagg atacaactag gtgggtctct 1800
 35 aactctctat gcagctgtgg tctgtggagc aaaacgatga aatggaagac gggacgttga 1860
 cgagggtgaa gaaaacgagc gtttgaccag cgtaaccai ggcgtgaaca gtagcaccac 1920
 taacctgacc gagaggttga agaagatgca atcaacgggg tactatagtt cccacgaatt 1980
 40 tcccagcaac aacgggttgg ttctcactac tcacgaattc cctgtggctc aacaactact 2040
 agtacaacct ttgtccatt atgataaaag ttctatctta atttttattt acacgttttt 2100
 caaactgttt ttaattttc tatataaaaa atacttaaaa tatcaaataa aaictatitt 2160
 45 tggagtittta aaaaactcaa tlaatcalat alattatga ctatttttat ttacgttga 2220
 ctaaaatate ttcacttca tttaggttat gtcttttct catcaagata catgatacat 2280
 50 tagcatgttt ttcaaactgt tttttaattt tgtatataaa ctactctaa aatatcaaat 2340
 aaaatttact tttaggggtt ataaaagtaa aactcaatta atcattacta acttgtttca 2400
 ttttactgtg actaaaatat ctcatcttc atctaagggt gtgtttggat ccaaggacta 2460

aatlttaatc cctatcacat cggatatttg acactaaita gaagtaitaa acatagatta 2520
 5 atgaigaaac ccattccata accciggact aattcgcgag acgaataiat tgagcataat 2580
 taatccatga tttagcctaig tgaigtctga gtaaacatgt actaattiacg gattaattaa 2640
 gcttaaaaaa ttatctttac gaattagctc tcaittatac aattaattit attgttagtt 2700
 10 ttagtttaat actltttaatt agtatacatc cgacgttaaca ctgactgata caaacaccaa 2760
 ctaaatcgaa aatcaccgaa tggctcgtca tccctccaca tgagatgcca agatggaaca 2820
 ccaacaatcc aacggctagg aagcgcccca tcccaccac cgcctaaccg ccttccatg 2880
 15 caagtgggtc ccaccccttc ctctctttt ttttctttt tacaatccc ctctcccttc 2940
 ttggctagct agctagcttg gcccaacgcc acgagccgag ccgagcacat ccgagaccaa 3000
 gccgagctca ggcctcagc tccccctct cctcgtccca tttccggttt cctcctccga 3060
 20 tttccccaa atccgcacgc ctctccctc cgcctccatt tttccgatt cccaattccc 3120
 aaatccggat cagccgcagc cgcagcagca aaaaatttcg aaatccaat ccaaaccat 3180
 cccccccag acgacgtcac ccacatccc acccccgga gacgagacga gacgacccc 3240
 25 aaatctctct ctctctctc ctatgcgcgc cgccgcgcgc gccgcagcag cagcagctag 3300
 gaggcggagc agcagcagca gcagcagctg agatgatcgt gcgcacctac ggccgcagat 3360
 30 cccgctcctt ctccgacggg ggaggagggg agcgcgcgcg cggcggggg tctcgtcgt 3420
 cgcaagacgc gtctgaattc gacggggagg aggaggacga cctcgtcctg ctggggctgt 3480
 cgtcgcagtc gtgcacccg cccgcgcgt cgcaggagtc gtctcgtatg tgggacttcg 3540
 35 acgaggaccg gccgcgcgc ccccgcgcg gccgggggag ggggtgggggt ggggactacg 3600
 cggagcccg cagggcgcg gcggcgcgcg cggcgccac ctgcctcatg gaggcggagg 3660
 agtacggcga gatgatggag agcgtggagc aggcgaactt cgcgtctgac gggctgcgcg 3720
 40 ccaccgcgc gaggcgggtg cgccgggcca gcttctctgc gctcctcggg atctgcgcct 3780
 ccgcgcgcgc ccgcgcgtc ctccgggccc aggggtcgtt acaccaaaga accttcttt 3840
 45 ttttttctt actgtctgc gctgtaagta aagaataaca attcgcgtc ttgtctttgc 3900
 ttgcggggca atcttggtga ggaatctgt tagggttatg aaattgggca gccagttctt 3960
 gtctctctg cgtaatctg gcggaaacag tgggattttg tacgattatg gctccgtaat 4020
 50 cggcatttct gtgggaaatg aaccacctt agggcatttg accttcgaac agcatgcttg 4080
 gtgttgcaat ccgtagctat tgccttctc ttaggcacia gaacttgctc tgaattatga 4140
 ttlaaccaact tgtgttgtt ttctgttct gagtttctt gcttggttag ggttagggtt 4200

atcaccgtgg tgggtgcagaa ttagaigtic gclactigic ttaaccicig ccttgcccaa 4260
 5 tttgggtaccg agtgttacag ctgggtttag gaagtgtgat ctttgagcat ttctagcatg 4320
 ttgggtctctt ttttttgcta atctcacatg gtgttagagg aaggaagcat agtgactgat 4380
 gatgaatgcc tagatactag aaatcacatc ttattaacig aattaggatt gcttgggtat 4440
 10 ctatgttagat atgactgtag aatgttacig ctggaaatgc tatccaatat ccatigatct 4500
 ctagcctaata atatctctcg aggccaagag atcagicaat ttigaacttt caggagagtt 4560
 tctatttgggt acttaatctc ttttatttgt tacttttgggt gccctggctct cttttcatga 4620
 15 ttgctaagta gacaggtaaa gtcttaccia aaattatct taaaagtcca aaatcgcttt 4680
 agattaagga gtcgacgcca gagccttagg cagagtctta taaacccaaa gcacaatgct 4740
 acaatgttca caaaactttt gtggaatttc cacttgagct gtataaacat cgcaatctac 4800
 20 tglgaataaa agaagcacit gatggaagtt catgttagca aatgacatgt ttctgtgtgag 4860
 gaggttgatt gcttgaacig ttaaggactc ttgcaacttt ttattttact tegtacccat 4920
 ttatgtctaata gtgcacaaat aaaatigctg agagtaaaaa tgtacaactt gttacgcacc 4980
 25 agcacacttc ctatttgtat ccattttcct gtggaatttc aaatgtattc aattgctgaa 5040
 atgtttccat tcaacaaaca catattccgt taatgaaatt attatacatt gcgttttgtt 5100
 30 ttcttactca caagtgtcct cttttcttat atcctataga ttgggtgcaac aaattattga 5160
 tgcaattttg gttttgaaca ttgatgatcc tcccigcact attgggtgcag ctgctcttct 5220
 attcgttttg gcaagtgaig tgagtlaccic tcaatcccat ccttgtgtct ctgtgcatgc 5280
 35 ttcatcttat tttttacgca tatcgattgt ttctttttat ataacagccc ataaaaataa 5340
 tcacatcaig gcaaagttaa ttatttctcc agtacagtta tataagtatt caccactttt 5400
 ccatgaatat ctggcatgt gatlacaaag aagattattt aagaaagtc atgtttttat 5460
 40 ttcatcattt tgtttgaagt tgaactttaa tttatgggtgt aaatttcagt taatatlgct 5520
 agcagctcgt attctttaat ggcatlaact cacttgtgct taltctccaa tatctccctt 5580
 cttgttgttc aggttcaaga aaatcatttg ttggaltcag aatcttgtgt ccattttctt 5640
 45 cttaaattat taaatccctc agtgaatctt gttgatcca aagcaccatc gataggttcc 5700
 aaacttcttg gaatcagtaa agttcaaatg cttaatggat caaataagga ttctgactgc 5760
 50 atticagagg aaatcctttc aaaagtigaa gagattctct taagctgtca agagatcaag 5820
 tcgctcgaca aagatgacaa gaaaacaaca aggccagaac tgtgtccaaa gtggcttgct 5880
 ttgttgacaa tggaaaaggc atgcttgctt gctgtttcag tggagggtaa gttttaatca 5940

aatttcttgg tcatgatitc cctttaigac cattataatt atttttatga gccaaataag 6000
cagttgccat aagttacata gcaccigtgt acaatattca tgggtgggtt gcttagccct 6060
5 ttgcttcacc tgcctttgat tgaigacttc catccgtgtt gcacaactga attgagtaa 6120
ttgactgcac tagaagcacc tatggccatt gtcatactag gaaggttttc ccttatcaaa 6180
10 tatitgatitg ttacagagac ttctgacact gtgtccagag tggaggaaa ttttaaagag 6240
acattaaagg agttgggagg tcttgatagt atttttgacg ttatgatgga ttgccattca 6300
acattggagg tgagatctcg ctaacatcgc atattttaca cticctttgt tcaactctaa 6360
15 aggaatgggtc aagttttgtt cctttttgcc attttagctt laaigtgctt gaagccacat 6420
gaaagcaatg ctgttcaga tacaatagcca aaggttgtta tattttggga catggaaaat 6480
gcttagaggta gtaactaatt tcatcaggac atggaaaatt ggctgcatca caaattatgt 6540
20 tgtttcatgt tgcaaaatag ttttttaata cttttttatt ctgcatgtgg tgttagtgc 6600
ttacagtgat tctctgag attataatcc ccacgataat aatacttgac atatctacac 6660
caagtggaca ttattcaatt ggatgttact ttccagcta tacttgctgt tcttgcataa 6720
25 actttggagt aaattgcgt tccctttaag agataaactg ctgggtgctc ctatctgtgt 6780
actttttaig ccccaacta ataatgcaat catattacgc tgataaactg aataaataaa 6840
30 ttaacaatat acttctgggtg gaaaccttgt gtaicagaat ctcataaagg atacctcaac 6900
ttcagctttg gacctaaatg aaggaacatc ttgcaaagt gccgtctctc tcttgaaatg 6960
tttgaaaata ttgaaaatg ccacatttct aagcgatgat aacaaggtaa tgttccattat 7020
35 atattctgtt tcagtttagt acccatttct ttctctgtta ccatcttctc ccttcatttg 7080
ttctgtgcaa aatgtgcaa cagtgtgact ttgtatttct gcttaacatt ttctttttt 7140
tcttgaaaag cagtataaac tcttacactc attttgcttc ttgcagacc atttgcttaa 7200
40 tatgagtaga aaattgtacc cgaaacgctc ctgctttct ttgtttgttg tcatlactag 7260
tattattgag ttattatcag gtatttttct taataatata atatgtccgc taacacaata 7320
aaatgtttta aacatccagt algttaaagt tgcagtctga cgcttatttt gttttgctgc 7380
45 agctctttca atactgcaga attcttctgt tgtttccagc tctacatatc cgaaatcgtc 7440
taaagtctct caacagagt gctctggtaa taacaaacac caaatttgtt tgatcaactc 7500
50 gtttggcttt ctgtgcactg ttcaatata gtttggtcgc cattcaagtc tcaactacaga 7560
tgttgaactt gacctgacac ggtggcacca atattataa aacgtacct gatattttta 7620
atatticatg ttccctgacc cagattatct tgttggttcc tcalataagt ttaattagt 7680

tcttcttga aactttgta tgcagcagat gtcattgggg gaacttcatt taatgatgga 7740
 aagcgcaaga actcgaagaa aaaaaacctt ttgtcgaacc agacacgcca tagttgctta 7800
 5 tcttcaaaat cagaagttc tcatattact atactttctg gtagtgatgc tggctgtca 7860
 cagaaggcat tcaattgtc tccatctata tcaagcaatg gggcatcaag tggttcatta 7920
 10 ggcgagagac atagcaatgg tgggtctttg aagttgaata taaaaaagga tctgggcaat 7980
 gcaaattcaa tttagaggctc aagtggttgg atttcaataa gagcgcacag ttctgatggg 8040
 aactccagag aaatggcaaa aagacgccgt ctatctgaaa atgtaatcac cgacagtggg 8100
 15 ggcggtgatg accttttgc ttttgatgat gtgatcagg agccttcaaa ttgggaactg 8160
 ctgggtccaa aaaagaaatc gcttcagaaa catcaagaca aatcaggaaa tggagtgtca 8220
 gtgcaagtc atgaaccaga ccaacctgaa gatcttaate agtcgggtac aacatctctt 8280
 20 tttagtgcta aagatgaatc cagtctttg gaagactgcc tcttggcatc agttaaggta 8340
 attaaatatg ttctctctg atctttctg ttctttctc aagagaatat acattcttgg 8400
 gtcacagttt ctgggtttgt cttgtgact ttgttgagt acatatttg aattcacaaa 8460
 25 atttctttt caatatggct cctcaatcta tagcatctgt cgtgtatgta ttctgtacaa 8520
 aatagtattg taacatctcc tagaagaaat tggcaccatc catacatac agtagcaatt 8580
 30 tatgagacgt gatcttgatt ggaggtttag gacagagcct cgagctaaat tgctattgta 8640
 ttgtatctac tatcttttag tacatgatat gtgttgggca ctctgtgtct gagtgtatg 8700
 agtgcttaag ttacatagt tcagctaaca tgcatagtta agacagttaa tgattaaatt 8760
 35 taagtgtaga aagaaggta tttcaaaaga tttttaagga caatataatt gtttcaccgg 8820
 gactcatgct tgtctgact gtgagcctaa tgttacctt acatgccctt acattgtcta 8880
 tttttatcg tttatgaga tcttccaaac aacttgatct gtcttaatgt tttttgcta 8940
 40 gctctttct tggatatctg gtaaattgtt aggccgaagt atgaactttg ccttattgtt 9000
 tcaaagaaaa tgaacaact cctggaaaag tctaattttg gtggccctt attttgtga 9060
 ccgtattggc acacatctaa ttctgtgtt cctttctggc aggttcttat gaacttagca 9120
 45 aatgacaacc catctgggtg tgaattgatt gcgtcatgtg gtggacttaa cacatggcc 9180
 tcttgatca tgaagcatt cccctcatt tgtttgtcg tggacaacaa ctataacacg 9240
 50 agagaigtca atcttgatca tgagttaaca tcttctcaaa acagcaaggc acaccaggtc 9300
 aaaattaaagc aattgcgaga tcatgaactt gatttctgg ttgcatatt gggcttgctt 9360
 gtttaacctg tagagaagga tagccttaat aggttaagtc ctacatgct tcttccatt 9420

1 tgc tcaattc atatcagigt tacigtictg gcagttccit ggggtcagga ctacagaaaca 9480
 2 tccaattaat gticattgtc tcttaacgac tcagaaatac ttataacct ctccacaggg 9540
 5 tacggcttcc atcigcccggt gticcigtg atctatcica gaatccacag agtgaagaga 9600
 6 cacagagaga tgcataagca ctctctgtt ctgtattctt agcaagtcaa ggtgctagtg 9660
 10 aagcttctgg aactataca ccggttaattc aaaattcttc aagttccttt tgtaigttaga 9720
 11 ttatactctt glaaaactcg gcatttatta cctgctcttt gttcaaaaa gcagtatttt 9780
 12 attttgcctc ttagcatagg tcagcagaac agttgatctt attcagaaaa caatattttg 9840
 15 catgtaacat actgttatct atgagatgaa aattaatgca tgtgtaataa tgcattgat 9900
 16 aaatatttgc tatctgaalc cagtctacca acctagttt gaccgaaatt actgagggtc 9960
 17 tatttcaaag aataatttag tgcaccattt gtccaactac tatgaagtaa aatggatttc 10020
 20 ccttctattg acatcgggtt agaagtgaag ggccatctta atgcaatgtt ctcaatgcca 10080
 21 caaaccacaca aatttcattt acacatacag attattatta acatagctat aaattggatt 10140
 22 tccagaagct tgaattgaat ttatttgggt acaattgaaa gcactgggaa cattagcatt 10200
 25 tttttttagt tcttgggtat tgcaatttat aatgttatc agaacttgtt acctcacaat 10260
 26 gcaattcatt tgcattctt tgaaccttt gatgactgt tgcattgaaa caacaggatg 10320
 30 atgaggagtc ttgatgcaa ggagcacggg aagctgaaat gatgatcgtt gaggcctatg 10380
 31 cagcccttct tcttgcgtt ctccaactg aaagggttgc aatctgtagt tgatggattg 10440
 32 ttttattaat gtctaactac ttgcataatg tcagcacat ggcatltaac ttatactgtc 10500
 35 tgttaactgc aacagcatga aggttcgtgg agccatttcc agctgccctc caaataacag 10560
 36 cttaaaaaac ctgtgcccgt cgctagagaa attgttggtt tgcctccata attcttgaac 10620
 37 tactgtttgt ataaaaaagt atggaatgc ttgaattta ctccatttgg gaaatcattt 10680
 40 atttttcatt tctgagggtt gagggttcac cataattgtt ctcccatcc aggaagccctg 10740
 41 ttgcaaaaat ttacataaaa taaggaaaaat ttgaacttgt ttcaagtttg aatagtaaca 10800
 42 ggaattttta ttctcaact ggagaaaaca ttccggctgg gacttttaac ccttaaaaatg 10860
 45 ctatgttgtt cccactgtta gatgtctgtc tgcacattt gaaactttgt gtaatacttt 10920
 46 tatcactacc ctgagatga gagacacaat ctggtaccga gtaagtta tgcatactcc 10980
 50 cagttgaagt acagcaccaa atcaagccaa catgttggct acgtaattaa atgttctctt 11040
 51 acaacagata gaggtaaaaa gggagtttct aagtatctaa cctcttacc tcttggctta 11100
 52 gcactccagg cacaactctt tcttaacttg cgatttagga ctgactctg agaataattgt 11160

gtgcccacac tggltgagtg catgcciatc taagcigcta gtttttggtc attttgatta 11220
 actctgaagc tgcctgagct tatictgctt ccatcattta ttaatccatc atgtttctct 11280
 5 ttcagtcgtt ccatctgcag ctcaalaiga tcacagagga aacgcactca gctgtcacag 11340
 aagttatcga gaaatgcaaa ctctcataga aagagtgaag aggggccigt acagatcaac 11400
 10 taacaacctc ttgcagcaa aaaagcalac acacaagigt ttgcttggc ctggggctct 11460
 gcagatggac tgatactctg acctgcagtg ggcttgggag ctaacaatgg ttcatctctt 11520
 ttttttttta tgttttcccc tgttgttttt gctcatgttt tgtgtaattt ttctctctca 11580
 15 tctagcgaag ttaattttct tagcatgatg ggagtagccc tctttttttt ttctctctaat 11640
 taagtgtaaa gttagcaacag catagggatg aatgttcagt gtagtgtgtg ggttttcagt 11700
 taticagaga cgtccataca gtgtgtacct tgtgaccaca cgtcttaatc tgatgaagct 11760
 20 tagaataaat cacatgttag caatgcaata tcatctgcgt ctctctctac ttgggtggcc 11820
 atcaaattct ggttagaagt gtaggttgg tgtgtgtgtg caaatgccgt attccgctct 11880
 gttttgtgga agttaagaag tccctagtgt aaataccgat ttctcatgat ctgggagatt 11940
 25 gatgcaactc tgattgcagc atttcttttt attagaatgt acactccatg ctatcatgat 12000
 gtttatgttt tagtactaca agatttgggt aaccattatt ttaataatcat aataatttta 12060
 30 taaaatcttg gagtaacaag ttcataatca atgtagcat aactttttga ggctagtcta 12120
 tgtatatgtt ctctttgttt tttaaactaa gcactcaata aattattgat ggctgttaatt 12180
 ttctgaaggt ttaccgggt tggcccggt ctttataaat agcttcggca caaaagacaa 12240
 35 aacgggccct ccaacacata aatgggtgag ttacgttttt cattatcttt ggtaaaatca 12300
 agtccaccac gtagacactc ataacaaaag ttgaataatc ctgagaaatt ttgacttgag 12360
 tctatctttac ctgtgatac ggacatccaa ccttccctcc ctccctgaac ttatatttat 12420
 40 tcatattaca cctgaacttt atattattca tattacacce tgaagtgggt ttcatttaat 12480
 tgcatacatg ctgaaatagt ttgacaacgt gagatgcaca aaatctacac gtctgtctta 12540
 agttgcaatt cattttatcc cttttctttt tctctcttac ataggaatat caatagtact 12600
 45 aattcacatt acaatatagt ataaatgggt gatcgattat tggcaatata ctatattaaa 12660
 tattcaaac tagtcattta agctgccaaa taagttaaacc actatcgaaa accacaatat 12720
 50 aaatggcaatt acaaaactta ggggggtgaa tatccaattt laaagttcat gatgctagag 12780
 gaatttctat caaaagttaa tgggtacata tggacttttt cttttttaaa agaagctatt 12840
 ctatctgtaa acgttaaata tttttgttac ttattttttt atgattgaaa aaaaaactta 12900

gttttcaaaa tgattggctt gtatacaagc atcaattaga cttaataaat tcaictaaca 12960
 5 gtttcctggc agaaactgta atttgTTTT gttattagac tacgtttatt atttcaaatg 13020
 tgtgtacgta tatccgatgt gacaaccaa cccaaaaatt tcccctaact ccatgaggcc 13080
 ttacagatat atttgatggg tgtaaagttt ttttaagttt ttgggtgcaa agttttttaa 13140
 10 gtatacggac acacatttga agtattaaat atagacaaat aacaaaacat attacatat 13200
 ctgcctgtaa acaacgagac aaatttatta agcctaatta atctgicatt agcaaacgtt 13260
 tactgcagca tcacattgtc aaatcatagc gtaattaggc tcaaaaatat tctctctgta 13320
 15 attlacatgc aaactgtgta attggTTTT ttttctgta acatttaata ctccatgcat 13380
 gtccaaatat ttgatgcgt ctttttggcc aaattttgtt ggaatctaaa caaggatcaa 13440
 atttgctgaa ttttccaga cgtcacggct tgttcatcca tegtctgcat cgcgattcgc 13500
 20 caccgacgcc ttggtttcca acgaatttta tcatccgttt aaatataccc aaagctctcc 13560
 atcgccatcg gcggccaacg gcgaccgtc cgtcttacc aatccacca tccactcgcc 13620
 gccgccccct gatccaaagc ctccgccg cgccgtcga gaggaggagg aggaggagga 13680
 25 ggaggaggag gaggcgtgag cccctatggg gacctcttc cgcccgctc cgtctgcca 13740
 cgccgccggc gccggcgacg ccacgccgtc gaccgcgcac ggtagccacg cgtctctga 13800
 gaggcccccc ccccgccgt cgcgatctc tctctctc ctgtttgggt ttgggtttgt 13860
 30 gatttgggtg tttttttt tccgcagcgg tgggtgtag cgggtggccg ggccgtggcg 13920
 tggagtgcca gccgcatcg gtgcgccgc gcccggttc gcaggttgcg gtggcgacgg 13980
 35 cgagctggag gaggcggagg gagaccgtg tgagatcgga ttctgccgt ggtggtgccg 14040
 ctaccatggg ggattcgccg caggcgtct caggtttga gctcttcca ctctctctc 14100
 gcaaaatgtg ttgctatgt cctctcgct ggctggcctc atagccatta atgtagttg 14160
 40 ctggaacatt acattcgga cgttgttggc aattgcttga caaaatgtg aattgtggag 14220
 gggagaaaaa tegtgtgaac ctgcagtgac aaaattgcca tctataatt taaaactgaa 14280
 ggtgtggaaa tcaaacata tcatggccag cacatcatt ttgttaacca cttgacata 14340
 45 ttgttggcct ataacagta gctccacacc aacttggag ggtcaalg aatgtaagta 14400
 taaattgagg ataactggca gtgttaaga cttctacag aactgtagc agctaaaact 14460
 50 agctattgtg catttatgt tcatggaaat tgagcggaaggatatt ttactaagac 14520
 gtataatgca aaacaaaaa aaaaaaact atgtctatgc agtttcatg taatgtgcgg 14580
 atgcaaataa aatcatgtt atggacaaac taatgggatt cataccaaat tccagaattg 14640

catttcttat gtggttactt ttgtttgttg atttggttac cagacatcga tgtggtttca 14700
 5 agggtcagag gggtttgctt ctacgcggtg actgcagttg cagcaatctt ttgtttgtc 14760
 gccatggttg tggttcatcc acttgtgctc ctatttgacc gataccggag gagagctcag 14820
 cactacattg caaagatttg ggcaactctg acaatttcca tgttctacaa gcttgacgtc 14880
 10 gagggaatgg agaacctgcc accgaatagt agccctgcig tctatgttgc gaaccatcag 14940
 agtttcttgg ataictatac ctttctaact ctaggaaggi gtttcaagtt tataagcaag 15000
 acaagtatat ttatgttccc aattaitgga tgggcaaigt atctcttagg agtaattcct 15060
 15 ttgcggcgta tggacagcag gagccagctg gtatggcigt agtctcatcc ctgccttctt 15120
 aagtagacat atatacattt acagtatttg gtaataaac aagattttat gaatcatata 15180
 tgattttggg gaaaacacaa aactctcttt gttaggtgcc ttgaacatag ttctgttcac 15240
 20 acagttatag caccctcttt aaaatgaaga acttgttgc atacacataa ggccaaacca 15300
 cataatgaat ttgtttatt tctatctttg aatgttagca tcgtttttgt ttaalgcatg 15360
 atcgcccttc tatataattg tagtaigtca acattgtatt ccatgctgag cataacaaat 15420
 25 ggtttgttaa aattcaggac tgtcttaaac ggtgtgtgga ttgggtgaaa aaaggagcat 15480
 ctgtattttt ctltccagag gggactagaa gcaaagatgg aaagctaggi gcatttaagg 15540
 30 tttagtaacc aaacttaggt tacattacat ctaatgagat ttttatattc agtatataat 15600
 gttaaccttc tcaiggigia ctgacgiggt tataaatgtc ccagagagg tgcaattcagt 15660
 gtggctacaa agaccggigc tctgtgata cctattactc ttctcgggac agggaaactg 15720
 35 atgcccctctg gaatggaagg catccittaat tcaggttcag taaagctcat taitcaccat 15780
 ccaattgaag ggaatgaigc tgagaaatta tgttcigaag caaggaaggi gatagctgac 15840
 actcttattc taaacgglla tggagigcac taaagaaaga tgggtttttt tttattata 15900
 40 tggaaacctat tcaaaggcac agacaggctt tcaaggctaa gcttgttaca ggtactgata 15960
 ctagtacta attactttcg taatcagiat aaataagctt gtgtagtgt atggcattgt 16020
 acatttctgc acttggtaaa ttacagaag aggcaaglaa tatittagag gattgagttt 16080
 45 attcaccag tcatatagtt gaagaggcaa gtaaccigia agagaggact gaacattaac 16140
 acctcttgtt cgattaaaaa tgaccaaaaga gcatcaaaca tgaattcgag gcigtactt 16200
 50 tagatatggc ccattaatlt gttagttgt ctatgtacat cctagtgtgt gtaaatgcca 16260
 gttaccattt ctatgaicta aaacaalcaa ctcttttagt atattttcaa aaacgaaaat 16320
 tcagtlacaca tgtatgaatc ttaatatctt tctctagctc gttacaaaag caacaaaggc 16380

accgtgtcag ctgggttcaca ttagctagtt tgtacttagc attatccact agcaccitat 16440
 5 tttcatgcat aicatgciaa ttigcttgcc cacgttaggt gggaattttt ttcattgttt 16500
 ataatttata tatgttttag acttctagtc cacaatttat gtacttcatg ttcctgagcc 16560
 tctagtatgg ctgatagcag actaggigct gagtgcigt cttttttgca gactgaagag 16620
 10 agaagaaata caagactgtc cattgttagt cagatttgta aaaatagact ctgatgtagt 16680
 ttacttttgc ccctatttta tttttaacaa tacaatatata taacagatcc taagaactta 16740
 tcttaattta ggagaagttg ctctgtttcat taaattaaat tgtgaagttaa aaatgtgtgc 16800
 15 tctgagtcgt caatgcaatc ctgtgttctt gtttgaagat atgggttagg gcaggccagg 16860
 atigaacact gaatggtaag actgccttcg ccttcagacg ttattgctaa attttttagct 16920
 actigcagtt agtgcgtcca cgccgattaa gcagttagaac aaagtatttt tgcgtgcac 16980
 20 aaatgagtta tatttcattg gaaatcgaag cgaaaacgaa tcaaaagtta gaagaaaagg 17040
 ggaaacttgg taattactcc ataaagagag tgcattttat tggtaagatg gtatccggaa 17100
 25 gctgtgagct ccgggctgta tgtattctgg caaatttgat atgagatgct cgattattgg 17160
 cttaagttag cgalatcaaa ttgggggaag caccaaagga attattgtga aggagttaig 17220
 ggtgcgtgac gttatctgct aggttcaaat ccttgtggct atgaatatit atctgctagg 17280
 30 ttcaaatcct agtgactatg aatatlaatg ggtaaggtaa gggattttatt gtttaattta 17340
 gtctctttta gatgtgcca tgggaccca ttggtaact gtaataatgc ttgtatttgg 17400
 attcacttgt gttacatgca cgcactaaac atgtgcttta ccttttcac tgtttttgcg 17460
 35 ttctgggcta gaaactcaaa cgttgaattt tccatggct gctcaacttg acaattactg 17520
 cgtgtcaagc gatcttatc gcatactatg cgcacaagt atgttatag gatatgatga 17580
 cagtataacg tgtgatattg atttttttta taaaaaatg atgttccttt ccttgatgaa 17640
 40 ggaacaaaga ctttttttta aagaagggtt ttactaaaaa caaaatgac aaaaacaaaa 17700
 tatcagtga catggcaagt gtgctcggca atttttctc tgtacttta acaaaaatac 17760
 45 ttctatatgt tcttttttat aagggtggca caaatcttt aaatgagcca aatatctaca 17820
 ttggatttat taaaaactgt ataaattata attatactc tgaaaggttg tgtgcatctc 17880
 tcttggagaa aatgtataag ttgcaacaa acattaatcc acgttatgta acittttttc 17940
 50 gccggaaagg ccgaaggagg cctgacggag cgtggggctc ctaccggga gaccgcgcag 18000
 gcccccttt gccggttcgg ccggggactc agggtgaaat tctaagctct ctgtatgtgg 18060
 aaggttcgcg accgtcgaag gagcataaga cacgggcgat gtatcacaggt tcgggccgct 18120

gagaagcgta ataccctact cctgtgtttt ggggggatct gtgtatgaag gagctacaaa 18180
 gtatgagcca gccctcccci tgttctgggt tccgaatctg gaaaagttca gtccagtcct 18240
 5 cccctctaag tgggcaaggt cctcccttta tctcttaagg ggataccaca tgcaccatct 18300
 cctccttttc tgtggggact taccctacct ttctataaat ggacggagat ttgtatagtt 18360
 10 gccgtccgaa tgaccttcig ataggacggc ccatacctac ctccacttcc gccgaaagca 18420
 ggtgcgacgt gggattatgg ctgtctgtct acgacaigac cagtgtcaga ctggtcacaa 18480
 attgtcatt cctgtccacc acgcgtcagt ttagcaatct acatgttggc ccttcttcac 18540
 15 acaacatctt gccgttaatg gttaggatga agcctggcat atatctaacc aggactaacg 18600
 tgccatctct aggaggtaac acgctagctc cagctgggga cgagcgctta gaagccctcg 18660
 tctgacggg atggggcgag gcgtgcgtca gatcgctgt cgcacctaa ccccgatct 18720
 20 gaccggctct tgaactgtca cagaccgat aaacgagtgc actgcacttc gttacatgcc 18780
 gcgtgacacg ctacgcaaaa ccgcaataaa tgggttagg tgagccccgc tgtgtcacc 18840
 taaccatac acgcggagca aaaaccacg aggggtcggg gcgcctcggc cctcggggcc 18900
 25 gaggcgggtg cggctcgacc cctcggggg gactaagagg agggcgaaca catcacctc 18960
 gggcccgacg tccccgagg gtgccaggcc acgtggcgga ttgtgtctgc ctcaaaccctc 19020
 tagtcatgat actcctgatc ccatgtcacc gacagtagcc cccggcgta tgcaggggcg 19080
 30 atcgccctct ttaagggag cggctcggcg tgacgccact cctaaggcct ggtgacaggt 19140
 gggaccggct tccacaattg ggcagaaacc caacggctac aaatcacgca catcggaat 19200
 35 ggtaactcta ctatcaataa tgagcggctt cttaagact gccacattac tgcagtagca 19260
 cacgaatctg gacatggcga ttctgttctg ctggagatat ggtaacgtcg ctltggctcg 19320
 cgagcgtaat taacgcgcgc acgataatga ctatctcgac tgccacaacc gcataatcac 19380
 40 ctcatgcgc gcaagcgggc gaatgggatt agtggaaagc tgggcgcgag aaacgagggg 19440
 gcgaaatagt gggcgcgaga agcgaggagc cgggcacagc gttggcaaga gtataaaggc 19500
 actgaggaaa ggatcgtt ccttcttct gccatcatt ccttgtctt cgcgcctgc 19560
 45 gccctaactc ctcttctct gtgtctact ttccgacac gcgtcgtct tcaatcttct 19620
 ctctctcgg cgccatggca cggggctcgg ctctgtctga tggtagcgtg ctgcgcctt 19680
 50 cccgcatctg gagcgagagg caggctgggc tggcgcccg ctcatgccg gaatctgcca 19740
 ccggccggga galatcacg ctgggtgagg gacgccggc gccagactac ccggggcgg 19800
 ccgtcttctt tctccccctt gcaatggcag ggcgtgtcc gccatttct tcttcttca 19860

1 tggatgttct gaagtcttac gatctccaga tggcgcacct ccccccaac gcggtgatga 19920
 2 catlggccat cttecgcat ctgigcgaga tglitcatlgg ggigcgccca tctcttcggc 19980
 5 3 tgttccggtg gttcttcacc gtgcagtcgg tctcgccgcc atcggtagtt ggtggctgct 20040
 4 acttccagcc atggggggcg gtgctgaatc gctacatccc ctgcgccctc cgcaagaagt 20100
 10 5 gggacgactg gaagagcgac tggttctaca cccccctgc cgacgaagcg cgcccttgac 20160
 6 ttcgagcca gccccggcg caggcctcca gctggcgggc gccggtagat ctgggggatg 20220
 7 gctatgacgc cgtccctgac cgcttggcgg gcctacgac ccaggggctc acaggggcca 20280
 15 8 tggigtacgg cgactacct cgtcgtcgga ttgcgcgct ccagcggcgc gctcggggcg 20340
 9 cctgggagta caccgggtcc gaagactaca tgaaggacca ccagggagtc agatgggact 20400
 10 gggctccctga ggatttcaag atagtggctc aacgggtgct gaatctcaac tccatggagg 20460
 20 11 cgtccctcat tccccagga atctccctc tctgcagcga tccagaccgc gcctccatcc 20520
 12 tgaccattat gacggcggtc ggggcctcag aggagtgagc tccaaagggc cagcagggcg 20580
 25 13 caggcgggag ccgtaggggg gatcaatcta cccgggagg gggctcgtgct tctgggtctc 20640
 14 gcgacggagg cccgaggagc agccgccctg ccgacgccg ggggaagagg aagcagggag 20700
 15 gaacacctcc cccatctctt ccccgagggg gcggggcggt gcgtgccaac agcaggcgcc 20760
 30 16 cggagggggc cgcgccgaca tgcagcccg agggggagcg caagaagaag cggctccgca 20820
 17 agatggggga gacagaacca tctcggggaa acctatttc cctccaaag tggctgttta 20880
 18 accgaccccc tgcaggltc gtctctcacc catcgtggct gtattcattc tctcaacgcg 20940
 35 19 agtttttact caccalcit gttegtctc tggcttttc ttctgtttca gcgagatccc 21000
 20 gtcgctccc tcccgccatt ccaagtcgg ccagcttgag gccgaggatc cggcggccgc 21060
 21 agaggcccg aggcgggaat ctgaccggcg agaggccgc gatcgctac gggaagccga 21120
 40 22 ggaggccgcc caggaggccg cccgggctcg ccaggcgag gaaaccgctc gggaggaggc 21180
 23 cgcccgggcc cgccaggccg aggaagccgc tcgggaggag gccgcccgag cccaccaggc 21240
 45 24 cgaggaagcc gctcgggaga aagccggtt tcgccaggac gaggcaatgg cgacttccga 21300
 25 ggcagctcgc gatgaggctc cgggcgcgtc gcttagagcc gcttctcgg gcgacgtca 21360
 26 ggcgacaact tccggggcag ctggcgacga ggcigcgggc gcgtcgcttg ggcccactcc 21420
 50 27 ctcaggcgac gcccaggacc aaccaggctt gagggacatc cccgagtcg gcacttccat 21480
 28 cggcggcccg agccgcgtgg catctctcc aaggcggctc tccccacgc ctctatcgc 21540
 29 cccgctgagc gcagagcccc ttctgcaggc ctggccgcc gcaaacatcg cgggtgtgga 21600

cgggcttagt gccaggtgg aggccctgca agcagagtgg gcggagctcg acgccgcgtg 21660
 ggcgcgltgc gaggaggggc ggcgctcagt ggaggccatg gtggaggtgg gccgcaaggc 21720
 5 acaccgccgg catgtctcgg agcttgaagc ccglaagaag gtgttggcgg aaatcgccaa 21780
 ggaagtggag gaggagcggg. gggctgccct catlgccacc agcgtgata acgaggcgca 21840
 10 ggacacccctc cgcccttcaat acgggagctg ggaggcggag ctagggaaaa agctcgacgc 21900
 cgcccagggg gtgcttgacg ttgccgtgc ccgagaacag cgggcggggg agaccgaagc 21960
 ggcgctccga cggcgcggaag agacccttga ggcgcgcgcc atggcgctgg aagagcgcg 22020
 15 ctgcgtcgtg gagagggatc tggcggaccg cgaggccgcc gtcactatcc gggaggcaac 22080
 actggcggcg cacgagtcgg cctgtgccga agaggagtc gcactccgcc tccacgagga 22140
 cgcgctcacc gagcgggagc gagctctcga ggaggccgag gccgcggcg 22200
 20 ggacagccctg tccctccgcg aggcagcgca ggaggagcag gcgcgccgca ctctggaatg 22260
 tglccgcgcc gagaggaccg cactaaacca gcgggccgti gacctcgagg cgcgggagaa 22320
 ggagctggac gcgaggcgcg gcagcggcgg ggcggtcgc ggcgaaaacg acttagccgc 22380
 25 ccgctcgtc gctgccgaac ataccatcgc cgatctgcag ggcacgctaa actcgtccgc 22440
 cggggaggic gaggccctcc gcttggcagg cgaggtaggg cccggcatgc ttgggacgc 22500
 30 cgtctccgc ctagatcgcg ccggtcggca ggtgggctc tggagagggc ggaccgtaaa 22560
 gtacgcgcc aaccatggag gccctgcccc gcgctctcgc aagatggccg gggctctcca 22620
 acggctcccc gaggagctcg agaagacaat taagtcatcc tcgagggacc tcgccaagg 22680
 35 agcggtaggag ctctgactgg cgagtlacca ggccaggac cccaattct ctccatggat 22740
 ggcgctggat gattccctc ctgggaccga ggacagcgcg cgcgaggc cgggatgccg 22800
 ccgaccatat cgtccacagc ttcgagggtc cagccccctg gctcgcgttc gcccccaact 22860
 40 ccgacgagga ggacaatgcc ggtgggtgcag acgacagta cgatgaggcc ggcgaccgg 22920
 gcgtatcgga ttgatcccc aagccccgc catcttcag tttttcttc ttttcttct 22980
 tctaaggcct tcgggctct ttttgtata gatcaacta atctgtaac aaaaatgaag 23040
 45 aaattttgt gtcaatttca tcttgctgtg tgtatgagat gaggatgac tgtacgtgg 23100
 tccitttgcg tcttagcttg attaagggtc cgtgccagg tcccagtcct caaaaggcgt 23160
 50 gggtcggggc tagtgccctg ggagatccac atgctgagac tggccaggcc gggaacgtgg 23220
 tgaccgaggg ttatgggtga ccgattgtg ggtttttgcc gattcccccc cggagttcac 23280
 cacgccccgg ggcacggctc ggctctgggc ccgctttggc gattttagcc gaccgagcc 23340

cccgagggca ggattgagca cgagtgacct atttcaagtc aagattcttc aaaaggaaaa 23400
 5 aaaaacacag atacagcctt taggaaatig aaactgcctt tatlgaaata ctgaaataag 23460
 agaaataaga atgtgcatgt ggggcagccc ccggccaacc ctgcacgccc gaggggggtgc 23520
 ggggttggcc cgagcccga acctgacacc cgaccccccc cctcaggggt agaagcgacg 23580
 10 aaggigtctg atgttccacg ggtaggcag ctcaatgccg tgcctcgigg ccagccgtat 23640
 ggagcccgcc cgggggacgc cgaccactcg atacggaccc tcccacattg gtgagagctt 23700
 gctcaatcca gcacgcgttt ggacgcggcg taggacgagg tcgtcgacgc agagtgatcg 23760
 15 ggcccgacg tgcgcctgat ggtagcgccg caggctctgc tggtagcgcg cggctctgag 23820
 ggccgcgcgt cgccttcgt ctccaagia gtgcagggtca tctctcgaa gctgatcttg 23880
 atcagcctcg cagtacaagg tggcccgagg agacctcagg gtgagctcgg atgggagaac 23940
 20 cgcttccgc cgttagacga ggaagaaagg cgtttcccg gtgtctcgcc ttgggtgtagt 24000
 tcggtttgcc cagagcaccg ctggcaactc ctcgatccat gaatcgccgt gcttcttgag 24060
 tatgttgaag gtcttgggtt taaggccttt gaggatttct gaattggcgc gctccacttg 24120
 25 gccattgctt ctgggggtgg caggtaggc gaagcagagc ttgatgcca tgccttcgca 24180
 gtagtcgccg aagagttcac tagtgaatig ggtgccatta tccgtaataa tacggtagg 24240
 30 cactccaaac cgggccgtga tgccttaat gaatttaagt gcggagtgt tatcgatctt 24300
 gacgaccgga taagcctcgg gccacttagt gaacttgctg atcgcgacat acagatactc 24360
 aaacccgccc gggggccgcc taaacggctc caggatatcg agccccaga cagcaaatgg 24420
 35 ccacgaaagt ggtatggct gcagggccctg ggccggctga tggatttgc ttggcgtggaa 24480
 ttgacacgct ctacatcgcc ggaccaggtc gaccgcata ttgagagctg tcggccaata 24540
 gaaacccctg cgaaaagctt taccaccaa ggtgcgcgag gcggagtggg ctccgcattc 24600
 40 gccctcatgg atatcgcaa gaagcacaac gccctgttcc cgaggaaatgc acttcaggag 24660
 gattccatta gccgcgcgcc gatagagggt ccttctacc agcacgtagc gtttgagat 24720
 gcgatggacg cgttcactc ctgcgggtc ctgggtaaa gcttatactg tgaggtatgc 24780
 45 ttggatctcg gcaatccaag caatcaatct aaggagctg ggagcgctcc cctcggttcc 24840
 cgaggcctgg acttcaacgg gccctggggg ccggtcaggc gcgtccgtct cccctaaggg 24900
 50 gtccgggtcg gccgacggct gggcaagcct tcttcaaag gcgccgggtg ggtcttggc 24960
 tcgcgtggac gcgagccgt agagtctgc ggcaatcatg ttatccgtc tgggcacatg 25020
 ccgaagctca atcccgtaa aatggcgctc catagccgt acttggcgca cgtaggcgctc 25080

catctgcggg tcagagcacc ggtactccit acagactlgg ttaacgacca gctgggagtc 25140
 5 gcctaacacc aggagggcgc ggatccccag tccagctgcc accttgagtc cggcaaggag 25200
 tccctcgtac tctgccatct tgttggctgc tcgaaagtcg aggcggacca agtatctgag 25260
 gacgtctccg ctgggagagg tcaacgtgac cccgcaccg gcgccctgaa gagacaggga 25320
 10 gccgtcgaac tgcattaccc agtgggcggg gtgaggcagc tgcgaggggt ccgtgctggc 25380
 ctgggggatt gagacgggt cgggagccgg ggtccacit gccacaaaat cggcgagagc 25440
 ctggctcttg atagcgtggc gtgggtcaaa gtgcaaatcg aactcagaaa gtctgattgc 25500
 15 ccatttcacc acccgctctg taccgtctcg attatgcaag atttgaccga gggggtaaga 25560
 cgtaaccaca gtgaccgat gcgcctggaa ataatggcgc agtttctctg aggccatcag 25620
 aatagcgtaa agcatctct gggcctgagg gtatcgggtt ttggcgctcc ggagggcctc 25680
 20 actaacaag tagacgggcc gctgcacct tgggtggggc cgatcctctt cgctaggggc 25740
 cgcatccctg gggcactctt cgtccaagca gcctcgggg gcgcacttgt ctctgtgtct 25800
 gatgacctcg gggctggagg ataacagggg cggccttccc acagtggctt tggggccgtc 25860
 25 ctgggggtca ggggtcctg gcgtcgtcgg acaagcgggc aaagggccaa ctccggtcgt 25920
 caggggcctt aggcctccgt tcggctcggg ggctctctt cctgtctctt tcccgggtcg 25980
 30 agtcagcaca gggttagctt cgggggtcaaa gggcgatagg tgcggcctc ccacagtggc 26040
 ctgagggcct tccgggggt cgggggtccc tagcacgtc tgacaagcgg gcagagggcc 26100
 aactccgtc gtcgggggcc tcgggccacc gtctggctcg ggggcctctc ctccctgtc 26160
 35 tctcccgggc caagtcggca cagggtgagg aagcgcgaaa tgagaatgt cctcatcgcg 26220
 ctccacaacc aatgccgcac taactacttg cggggctgcc gctlaagtag gtagcaagg 26280
 ctgtcttggt tccggggcga ccagaactgg gggagagctt agatagcct tcaactgggt 26340
 40 gagggcattt tcagcttct tctccagggt aaacggctcg gagcgttga gaagcttaaa 26400
 taagggtaac gccttctct ccagcctcga tatgaaccga cttagggcgg ccatgcaacc 26460
 45 ggtgacgtat tgcacatccc taagtgtgt gggggggcgc atccgtctta tagccgtat 26520
 ctctcggggg ttggccicaa tgcgccgggc agagaccaag aacccgagaa gcttgccgc 26580
 aggtacaccg aacacacact tatcgggggt taatttatg cgggcggagc ggagactctc 26640
 50 aaaagtctc gctagatcta tgagtaacgt ttcctgggtg cgcgtcttta caaccaagtc 26700
 atcgacataa gcttcaatat tacgtcttaa ttggctaccc aaagaaatc gagtagtacg 26760
 ttgaaaagta ggaccigcat tctttaaccc gaagggcatt gtcgtataac aatagggtcc 26820

1 atgggggta atgaacgcag ttttttctc atcttcccta gccatgcgaa tctgatggta 26880
 5 accagagtat gcatctagaa aacacaaaag gtgcgacccc gcagtggagt cgacaatctg 26940
 atctatgcga ggcaggggtt aaggatcctt aggacatgcc ttgttaaggt cggltgtagtc 27000
 gatgcacatc cgaagcttgc cgttcgcctt gggaacgacc accgggttcg ctagccactc 27060
 10 ggcgggggtt acgctgccat catattttc ggcgatggig ggccggaacc ttggggggcca 27120
 acggacattc cgaagactcg ccacaaaggc tctacagccg acaccaccaa ccgggggcac 27180
 ggagggctga tccccgcgc cgtgttgagg tgacactctg gacgaggaag cgccctccgt 27240
 15 tgcgtgggca gcacttcggt cattacgccg gcgctcgat ctgggtcggg cgtccggccc 27300
 cccacgcaga tctttctggg tcgaaggagt cgacgaagga gtggcgggccg aatggcgaac 27360
 agcggctgcc gctcgtcgtg cctccgctt tgacgacgcg gagccggttg tagcagcacc 27420
 20 agaggccttg gtggcggagg accgcccacc agcatctagg cgctgccgta ccgtcatgac 27480
 taatttggcc acgtcgicca gccatcgttg ggcgtggagac tccgggtcag ggacgacagg 27540
 cgggtgacgt aagagcgcg cgcgagcttg gagcgcgccc tggggcggtg tgcgctgcc 27600
 25 gtagacgagg aggcgacgt ccccatctcg ccgttctct ccatcgcccg cgatcggtga 27660
 agtcgcggat ctctcgaccc tctcgagcg cccccccgc ttaggacitt ggcgltggagg 27720
 30 gagcggltga gtacgagctc gacggcggtg gttcggctcc ccgtcgtcgc cactcacact 27780
 cggagagagg tctgtgcct ttgcltgc cgcacacagg ctgaacagga aaagcttggc 27840
 gcacacggaa gactacgaga gctcagaaaa acacacacig agtcccciac ctggcgcgcc 27900
 35 agatgacgga gcgtggggct cctcaccggg agaccgcgca ggccccctt tgcgggttcg 27960
 gccggggact cagggigaaa ttctaagctc tctgtatgtg gaaggltcgc gaccgtcga 28020
 agagcataag acacgggcga tgtatacagg ttccggcgcg tgagaagcgt aataccctac 28080
 40 tctgtgttt tggggggatc tgtgtatgaa ggagctacaa agtatgagcc agctctccc 28140
 ttgttctggg ttccgaatct ggaaaagtc agtccagtc cccctctaa glgggcaagg 28200
 tcttctttt atatcttaag gggataccac atgcaccatc tccctcttt ctgtggggac 28260
 45 ttaccctacc tttcataaa tggacggaga ttgttatagt tgcgtccga atgacctct 28320
 gataggacgg cccataccta cctccacttc cgccgaaagc aggtgcgacg tgggattatg 28380
 50 gctgtcgtc gacgacatga ccagltcag actggctaca aatgtctcat tctgtccac 28440
 cagcgctcag tttagcaatc tacaigtgg ccttcttca cacaacatct tgcctgtaat 28500
 ggtaggatg aagcctggca tatalctaac caggactaac gtcctatct taggaggtaa 28560

cagcctagct ccagctgggg acgagcgcccl agaagccctc gtcctgacgg gatggggcga 28620
 ggctgctgc agatcgccig tcgccacctt acccgcgatc tgaccggctc gtgactggtc 28680
 5 acagaccgga taaacgagtg cactgcactt cgttacatgc ggctgacac gctcagccaa 28740
 accgcaataa atgigggttag gtgagccccg ctgctgctac ctaacccata cacgcggagc 28800
 10 aaaaaccac gaggggtcgg ggcgccctcg cctcggggc cgaggcgggt gcggtccgac 28860
 cccctcgggg ggactaagag gagggcgaaac acatcaccccl cgggcccgcac gicccccgag 28920
 ggtgccaggc cagctgggag attgtgtctg cctcaaacct ctagtcaiga tactccctgat 28980
 15 cccatgtcat cgacaaggcc atccgaatgt attaaggagt aaaagttaca agaaaaaaca 29040
 ccacaatgca ccaaggtgca tgaccacaca ccatacacta cccccaagca caaaccactg 29100
 aggggtgaagc ctagcaccaa acgaccgcca ctaagtgtga ccaaacgccc ctaggccctac 29160
 20 ggagcaaca catagatgag acttcgaaaa cgatgccacc aagggtggta cgacatgtag 29220
 gatgctgcca tcttccatct aaaaagatgt ggctttcacc cagagaaact catcaagaag 29280
 ggagaggggt aaccttgac agcgcccca ggaggttacg acgcccgaag gcgtagccgc 29340
 25 tgccggctcg gtgaaccacc ggactaggct tccgcctagg acctatagc ctgatcgca 29400
 gatcaccgtc caccactcag aaccaccaca cagacaaaag gtacacgta gcttccaccg 29460
 caccgcaccg acgcccctc gtcggccgac tccatcgaac caccatccct gagagctggc 29520
 30 ccaggacccc tccgttccac caccgcccgg ccgccttgcc agttttggcc aaaggagaac 29580
 ccgggactgg gtgacattg ttcggcagcc tgagcttccc ccgctggcga gctgctgctc 29640
 35 caatccaacc tagaaactcc ccgcaaaaga aggggatgag ctclaggaag ggcgaggggtg 29700
 ccgaccggca acgaggaaga caaccatcg actccagctc cctttgact accatctggg 29760
 cctgcgcca tgccggatc gctgtcgtc cggctccggc gccaccacc tgcacccct 29820
 40 ttgcttggtc tccgcgccc tccggctgc gtcgcgccc ccagctggc gctaagggca 29880
 ccacgacggc cggccggcta ccgaggcctg gccgcgcca gggacagctc gcgctggcac 29940
 cagcgagcca cggccgtcgc gctgttgccg gcgccagcga gcacaaccgc cagctccaag 30000
 45 ggccgagcat gccactgagc cggccggct gccgcccggg ccggctgcac gtcaccggcg 30060
 cacacgaccg cagccgcca cgtccgct cgcgcccga ggcagccca tgccattgcc 30120
 50 gcgcacctcg cccgcccgt gccgagccgc caccgcgcac ctgctgagc cgccaccgcc 30180
 gtcctlagcc gctcgtgcc gccgccagc cagatccagg cgcgggatgg ccggatccgg 30240
 ccttgggggc gccggatccg ccgcccctcc acaccgccac ggctcacca cctccgaccg 30300

cagtgagggc ttcgtcgttt gcccacccct catcgcgctg aggaggaaga cgccaagaaa 30360
 aaagggccct gccgctgcc tccctgctcg ctgccggctt cgccgccggc gagctccggc 30420
 5 ggcggcgagg tgggggagaa gaagtgggga gggggcagct agggtttttt cgcccccaa 30480
 gccgcccglg cgagagcgac ggtggggggg gggggacttt ccaacctctt ccagtgttct 30540
 10 agttctccac gttatglaac tcaatttgtt taaccataga aagtaagaaa cctaccagcg 30600
 tgttaagctc tctttcattt cctttctctt tccgggtttt gcttccatca catgtcaagt 30660
 gaagggttct taactaccat tactcctaca calctaatit tttctcaga tctttcgag 30720
 15 gtatatatig atgctacatt ttatgatctt aagataatct ccttcacatt acctctgct 30780
 gaaactttag ctggaaccgt catcttcacc acaattigag cccaatttgc acagagcaca 30840
 acgagcaata gcttgccttt acgttcatta tttagcatga actactacta actaccaag 30900
 20 aatcaataca ccggtttaat aacgccattt talcacgta atatatgitt catccaacac 30960
 accggttttg gcacagttgc aaacttgcaa taaattcttt cctacttctc catccataa 31020
 tataacaaat tggatatgct cgtctggtac taagtigcta tattatgaga tggagggagc 31080
 25 acttcttttc tccaaaaa taagaatata gtattggatt agatattalc tagattcacg 31140
 aattcgatta ggttgcttag atttatagtt gtatgtaalg tataattcgg taataggta 31200
 ttacctctcg ggaaggagg agtagtttg actttttttt tctttataaa tgcctttgat 31260
 30 ttttatatta gtcaaatit atcgagttta actaagtlla tagaaaaaaa ttagcaacat 31320
 ttaagcacca cactagtttc attaaattta gcatggaata tattttgata atatatitgt 31380
 35 tctgtgttaa aaatgcigct atatttttct ataaacgtag tcaaatttaa ataagttaga 31440
 ctaaaaaaa tcaaacgac ttataatatg aaatggagga agtagtagac tataacaaat 31500
 ttaaaccgtg ctttgatttt agagcatcac taatatgta gcaataatct atccctaaaa 31560
 40 tttatttttt ttcctaaact gaaaatagga agtggaata ctctccatc taagagagag 31620
 cctaaattca ataaaaaact aaaaaactaa aggtggatcc ctclattaaa ctaccgcaa 31680
 aaatttatgt ttttttctc tccacgcgc gcagaacaga tatctgaltc aagttagcat 31740
 45 glaaaatttt taaagagata cttatagca ctcttccgt atttccaaaa gcaaaggat 31800
 ttaaaatcig actcaataa agatctatat atccaattta catgacacat gtttcgccga 31860
 50 atttttatat taataalaat taatatitit aaaattaaat tattagcaat ttgtttggag 31920
 gatttatcaa aacaggatgg acgttgttla taacagctc tagacctaga cgcgcttgca 31980
 aactgcggcc accttttat cacacaaatt ttgacaatt tgacacttc caaaaattaa 32040

1 ttttataaat taaccgtgac caaaacttat ttaaaaaataa tctttttgtt gagcgcaaaa 32100
 2 tcgtataact cagcgccaaa tagcacggcg ccgacctccc ccttccccct cccctctatcc 32160
 5 tccactgcig ccgcccacct ctccgtatca gctgcgtcgc gtigggttcc gccggcgctg 32220
 ctgctgcigc accagtcgc tagggcgggc gggcaltggc cgccgcgccg ctccccgcgt 32280
 10 ccgcgccggc gctgttggcg ccttctgcct ggagggtcgc acccaagggc gagggggccg 32340
 cacggggggc agtggcgccg aggacgcacg ccacgtgtc gacgaattgc tccggcgtgg 32400
 caggggcgcc tcgatctac gcttgaactg cgccttcgcc gacgtcgcgc gtcacagccc 32460
 15 cgcggccgcc gtgtcccgct acaaccgat ggcccagacc ggccgcgacg aggttaactcc 32520
 caacttgtc acctacggca ttctcatcgg ttctgtctgc tgcgcgggcc gcttggacct 32580
 cggtttcgcg gccttgggca atgtcattaa gaagggattt agagtggacg ccatcgccctt 32640
 20 cactcctctg ctcaagggcc tctgtctga caagaggacg agcgacgcaa tggacatagt 32700
 gctccgcaga atgaccacgc ttggctgcat accaaatgtc ttctctaca atattcttct 32760
 caaggggcig tigtatgaga acagaagcca agaagctct gagctgtctc aaatgatgcc 32820
 25 tgaatgatga ggtgactgcc cacctgatgt ggtgtcgtat accacigtca tcaatggctt 32880
 ctccaaggag ggggaltctg acaaagctta cggtacatac catgaaatgc tggaccgggg 32940
 gattttaacca aatgttgtta cctacagctc tattattgct gcgttatgca aggctcaagc 33000
 30 tatggacaaa gccatggagg tactlaccag catgglttaag aatgggtgca tgcctaattg 33060
 caggacgtat aatagtatcg tgcaltggta ttgtcttca gggcagccga aagaggctat 33120
 35 tggatttctc aaaaagaatgc acagtatgg tgtcgaacca gatgttgtta ctataactc 33180
 gctcatggat tatctttgca agaacggaag atgcacggaa gctagaaaga tgttcgattc 33240
 tatgaccaag aggggcctaa agcctgaaat tactacctat ggtaccttgc ttcaggggta 33300
 40 tgctacaaa ggagcccttg ttgagatgca tggctcttgg gatttgaagg tacgaaacgg 33360
 tatccacctt aatcattatg ttttcagcat tctaatatgt gcatacgcta aacaaggga 33420
 agtagatcag gcaatgcttg tgttcagcaa aatgaggcag caaggattga atccggatc 33480
 45 agtgacctat ggaacagtta taggcatact ttgcaagtca ggcagagtac aagatgctat 33540
 gcgttatitt gagcagaatga tcgatgaaag actaagccct ggcaacattg tttataactc 33600
 50 cctaaticat agtctctgta tctttgacaa atgggacaag gctaaagagt taattcttga 33660
 aatgttggat cgaggcatct gtctggacac tattttcttt aattcaataa ttgacagtca 33720
 ttgcaaagaa gggagggtta tagaatctga aaaactcttt gacctgatgg tacgtattgg 33780

5 tgigaagccc gatatacatta cglacagtac tctcatcgat ggataatgct tggcaggtaa 33840
 gatggatgaa gcaacgaagt tacttgccag catggctca gttggaalga aacctgattg 33900
 10 tgttacatat aataclitga ttaatggcta ciglaaaaatt agcaggatgg aagatgcgtt 33960
 agttcttttt agggagatgg agagcagtgg tgltagtcct gatattatta cgtataalat 34020
 15 aattctgcaa ggtttatitc aaaccagaag aacigctgct gcaaaagaac tctatgtcgg 34080
 gattaccgaa agtggaaacgc agcttgaact tagcacatc aacataatcc ttcatgggct 34140
 ttgcaaaaac aatctcacig acgaggcac tCGaatgttt cagaacctat gtttgacgga 34200
 20 tttacagctg gagaclagga ctittaacat taigattggt gcattgctta aagttggcag 34260
 aaatgatgaa gccaaaggatt tgtttgcagc tctctcggct aacggtttag tgcagatgt 34320
 taggacctac agtttaatgg cagaaaaatc tatagagcag gggttgctag aagaattgga 34380
 25 tgatctatit ctltcaatgg aggagaatgg ctglactgcc aactcccga tgcataatc 34440
 caltgttagg aaactgttac agaggggtga tataaccagg gctggcacit accgttcat 34500
 30 gattgatgag aagcacttct cctcgaagc atccactgct tcttgttt tagatcttt 34560
 gctggggga aaatacaag aatatcatag gtctctccct gaaaaatata agtccittat 34620
 agaatctttg agctgctgaa gccitttgca gcttgaat tctgtgttg agttctttc 34680
 35 tcttacagtt gtatttagagg agggatctt tctttatgtg taaatagcga gglatgtatg 34740
 tcacctctcc gaattatit tctctgggt cctagacggt aaacaagcaa ttatgttctg 34800
 ccttgatgc cagaaaaaac acaaaagtt gtcttatct ctactaacgg atcataaagg 34860
 40 aatttgtaac tggagtttca aacttaatt gtctaggcag tagttttggc attagatcca 34920
 acattgtgta ggattcatit gtgtgtatca atctataggg ttcatataa ttctgttaat 34980
 gtgtactgt taggtgttga atagtttgac ttgttttta actgaacaaa agatactgaa 35040
 45 atcgttccat tcaacaaaca catgttccgt taatgaaatt attgtacgt accitttgtt 35100
 ttcttactca caagtgtct ctltcttat atctataga ttgglaaac aaattatiga 35160
 ttcaattttg gttttgaaca ttgatgatcc tccctgcact attgggtcag ctgtcttct 35220
 50 attcatittg tgaagtgat tgagtaccct tcaatcccat ccttatgct ctgtgcatgc 35280
 ttcatlcaa tttttlacgc atatcgatt ttcttttta tataacagtc cataaagata 35340
 atcacatcat gacaaagtta ttatttcta caglatagtt atataagtat tcaccagttt 35400
 tccatgaata ttttgcatg tgattacaaa gaagattatt tgagaaaatc catgtttta 35460
 55 ttcatcttt ttgttgaag ttgaacttta atttatggig taaatttcag ttattatgc 35520

tagcagctcg taccttttaa tggataaact tcacttgigc ttatttccca atatctccct 35580
 tcttggtgtt caggttcaag aaaatcattt gttaggttca gaacttggtg tccattttct 35640
 5 tctlaaatta ttaaatcttc caglgaaact tgttgattcc aaagcaccat cgaatagggtc 35700
 caaacctctt ggaatcagta aagttcaaat gcttaatgga tcaaataagg attctgacig 35760
 10 catctcagag gaaatccttt caaaagtiga agagattctc ttaagctgic aagtatcaa 35820
 gtcgctcgac aaagatgaca agaaaacaac aaggccagaa ctgtgtccaa agtggcttgc 35880
 tttgttgaca atggaaaatg catgcttgc tgcgtttca gtagagggtt agttttaatc 35940
 15 aaatttcttg gtcatgattt ccccttatga ccattatatt tatttatatg agccaaataa 36000
 gcagtgttca actgtcata agttacatag cacctatttg caatattcat gggtaggttg 36060
 cttagccctt ttcttcacct gcttttgatt gatgacttcc atctgtgttg cagaattgaa 36120
 20 tiggagtagt ggactgcact agaagcacct atggccattg tcatactagg aaggttttcc 36180
 ctatcaaat atttgattgt tacagagact tctgacacag tctccagagt tggaggaaat 36240
 tttaaagaga cattaaagga gatgggaggt ctgtatagta tttttgacgt tatggtagat 36300
 25 tttcattcaa catiggaggt gagatctgc taacatcgca tattttacat ttcccttggt 36360
 caactctaata ggattgtgca ggcttgttcc ttttcgcca tttagcttta atgtgcttga 36420
 30 agccacatga aaglaatgt tgtccagata catagccaaa ggttggtata ttttggggca 36480
 tggaaaatgc ttgaggtagt aactattttc atcaggacat ggaaaattgg ctgcaacaca 36540
 aattatgttg ttttatgttg caaaaatagt tttttaatac tttttattc tgcattgtgt 36600
 35 gttagtatct tacagttcct ctgatgatta tatccccac gataataaca cttgaaacga 36660
 taataacact tgacatact acaccaagtg aacattattc atttggtatg tacttttcca 36720
 gctataacttg ctgttcttgc atgtglaagc aagtltggag taaattgcgc attaatftaa 36780
 40 atgcttgggtg ttccatctg tgtactttt attccccaac taataatgca atcatattac 36840
 gctgataaac tgaataaata aattaacaat alacttctgg tggcaaacct tgtgtatcag 36900
 aatctcataa aggatacatc cacttcagct ttggaccgaa atgaaggaa acctttgcaa 36960
 45 agtgcgtctc tcccttgaat atgtttgaaa atattggaaa atgcatatt tctaagcga 37020
 gataacaagg taatgtcct tatagttct gtttcagtt agtaccatt tcccttctct 37080
 50 gtactatctt ctctccgat ttgtctgtg caaaatgtc aaacagtgcg actttgtatg 37140
 tctgttaaac aattttctt tcttccgaa aaagcaatat gaactcttac attcatttg 37200
 ctcttgcag acctattgc ttaatatgag tagaaaattg aaccgaaac gctccttgc 37260

1 ttttttgggt ggtgtcatta tcaatactat tgagttatla tcagglatatt tttttaataa 37320
 5 tacaatgigt tgcctaacac aataaaatgt ttaaaacatc cagtaigtta aagttgcagt 37380
 ctgacgccta ttttggtttg ctgcagctct tccaatactt cagaattctt ctgttggttc 37440
 cagctctaca tatccgaaat cgtctaaagt ctctcaacag agttactctg glaataacaa 37500
 10 acaccaatit tgtttgatca gttagctctg ttggcttttc tatgcactgt ctcaatatag 37560
 tttgggtcgcc attcaagctt cactacagat gttagaacttg gccctgacacc aaatatitit 37620
 aaaatgctac ctgataatit taataatitca tgtttcciga cccagatitit ctgtttgggt 37680
 15 cctcgtataa gtttaattag tgacattctt gaagctttgt tatgcagcag atgtcatggg 37740
 gggaactitca tttaaatgt gaaagagcaa gaactcgaaa aaaaaaaact ttgttcgaac 37800
 cagacacgtc attgttgctt atcttcaaaa tcagaagtitt ctcatattac tatatcttct 37860
 20 ggtagtgatg ctgggtctgt acagaaggca ttcaattgtt ctccattitit atcaagcaat 37920
 ggggcatcaa gtggttcatit aggcgagagg cacagcaatg gttagtgggtt gaagttagaat 37980
 ataaaaaagg atcgtggcaa tgcaaatcca attagaggct caactggatg gatttcaata 38040
 25 agagcgcaca gtcttgatgg gaactccaga gaaatggcaa aaagactccg tctatcttaa 38100
 aatglaatca ccgacagtgg tgggtgtgat gaccttttg catttgaccg ccgcgtcggc 38160
 30 gtgccacca cgtaatcgcc cagctcgtg ccccgctgc cagctcgtg accgcgcacg 38220
 glaatacacac gcatctcgag gccgcgcta gctgatactt tctcatccgg ttgatttggtg 38280
 attttggcgt ttttgcatg gtgatggcgg ggggcgaccg tggccgaggc gtggagtggc 38340
 35 atccgcatca ggggtlatcg gccgcgctg tccgcccgg tccgcaggct ttggcggcga 38400
 gctggcggcg gagggagact gtggtgagat cggatttcgc cgttggtgtt gtcgtacca 38460
 tgggggattc gccgcaggcg ctctcagggt tgcagcctcc tccactctct tccctttttt 38520
 40 attttttttt ctgcataaat gtgttgat gtctgtctg ctgggcaggc ctcatagcca 38580
 ttaatgtagt ttgttggaac attacattt ggaacgttgt ttgcaattgc ttgacataat 38640
 gtggaattgt ggaggggaga aaaatcattt gaactgcag tgacataat gccatctcta 38700
 45 atttttaaac tgaagggtg gaaatcaaac ataalcattt ccagcgcac atcttgtta 38760
 accaccatga tatattgttg gttataacag ttagctccac accaaccitg aagggttcaa 38820
 50 tagaatgttt agtataaat gaggagaaca ggcagtgtt aagactttct aaagaacttg 38880
 tagcagctaa tactagctat tgtgcatit tgtttcatgg aatttgagca gcaatggata 38940
 tttcttacta agatgtatga tgcaatacaa aaaactatgt ctatcacgtt tacatglaat 39000

gtgcggaigc aaataaaaac atgtacatgg acaaactcat gggattcata ccgaattcca 39060
 gaattgcatt tcctatgtgg tiacttttgt tgttgalltg gttaccagac atcgaigtga 39120
 5 tttcaagggt cagaggggtt tgcctctacg cggtagctgc agttgcagca atctttttgt 39180
 ttgtcgccat ggtttgtggt cateccattg tgcctctatt tgaccgatac cggaggagag 39240
 10 ttcaggaaaa aaatttgaaa ataccatttt ttigaaaaag atttacgttt atatacacta 39300
 gtatgaagaa tttagcaaaa tataactaat ccgcagatcg gttatgcggg agcgcaacaa 39360
 aagtatggcg tggcggcgcg gagtggacgg ccgaggcggt cgcgcggaat ggggctgcgg 39420
 15 gaccgagcca gtctcgcttg ccgtaacgc ggaaccggta cgctcccgca gcgccagtgt 39480
 gcggaaccgc ggccccaaca ttttttact gcatggcact gtgtttaata ctgtttgaca 39540
 ctgtttctgg tactgtttta cacagttccc gggtcagttc cgcacaatgg aggcgcggca 39600
 20 ccgaccatga acaatgtgtg aacagtgtg cacagggtta aaacagtgt taaactgcgc 39660
 tgcacagtgc tggagtgcgt ggccactgcg gticcgcgtt ttggaaccgc gggaccgtcg 39720
 cgattccgcg ttttggagct gccggaccat gacggttccg cgcaggatcg tgggtccgt 39780
 25 attttgaatc tgcggaaccg tcgtgtccc gcgtttccgt ttgcgggat gcgtatatit 39840
 ttataaaacc tctccatgca tgtatataaa cataaattat tgaaaaaata agtatatttg 39900
 30 caaatTTTT tgcagagctc agcactacat tgcaaagatt tgggcaactc tgacaatttc 39960
 catgttctac aagcttgacg tgcagggaat ggagaacctg ccaccgaata gtagccctgc 40020
 tatctatgtt gcgaaccatc agagtTTTT ggatatctat acccttctaa ctctaggaag 40080
 35 gtgtttcaag ttataagca agacaagtat atttatgttc cgaattatt gatgggcaat 40140
 gtatctctta ggagtaatic cttagcggcg tatggacagc aggagccagc tggtatggct 40200
 gtagtctcat ccttgccttc ttaagtagac atatatgcaa ttacagaatt tggtaaaca 40260
 40 acaagatttt atgaalcata tatgatttg gggaaaacac caaactctct ttggtggctg 40320
 ccttgaacat agtctatc acacagtat agcacctct ttaaaatgaa gaactttgtt 40380
 gcatacacat atggccaaac cacataalga attttgttta ttctatctt tgaatgttag 40440
 45 cacttatatt tcatgcatat catgctaat tgcctgcca cgttagagtg gaatttttt 40500
 ccatgtttta taattatat atgttctaga ctctagtc acaatttate tacttcatgt 40560
 50 tcttagacct ctagtatggc tggtagcaga ctagggtctg agtgcgtcc atttttgcag 40620
 actgaagaga ggagaaatc aggacgtcc gttgttagtc agatttglaa aaatagactc 40680
 tgaatgtatt tattttagcc cctattttat atttaacaat acaaatatat aacgtatcct 40740

aagaacttat cgtaatttag gagaagtigc tegtittcatt aaattaaact gigaagtaaa 40800
 5 aatgigtgct cgagictgic aatgcaatcc tigtittctig ttigaagata tggigttaggg 40860
 caggctagga tgaacacig aatggtaaga ctgcttcigc cttcattigt gcacttggtg 40920
 cigccacgcc gattlaagcag tagaacaaag taattttgic gtgcacaaat gattlatatt 40980
 10 tcattgaaaa tgaagtga aatgaaccaa aagatagaag aaaaggggaa acttggtaat 41040
 tatalactcc acaaaattat tggtaagatt tgatattaga cgctcgatta ctgggcctaa 41100
 gtttaaggata tcaaaattgg ggaagcacca aaggaattat tigmaaggag ttgtgggtgc 41160
 15 ataacgttat ctactagggt caaatcctag tgactatgaa tattaatgag taaggtaagg 41220
 gatttatigt taattttagt ttctttaaga ttgtgtccgg gtacaccatt cggtaagigt 41280
 aataatgttt tgtattggat tcaattigt taccgtgcalg tgatttacct ttccatttgt 41340
 20 ttctgcgttc tgggtatgaa ttgtacgaga ttccaaggic agctcaacat atcagttact 41400
 gcgtgicagc cgaacttata tggtaagcgc acaagcgatt gtaacggat atgacagtat 41460
 25 aacgttgtat attgalacga ttttctttc cttaataaag gaacaaagac ttttttaaaa 41520
 aaaagaaggg gtattactaa aaacaaaaat gtcaaaaaca aaataacagt gcacatggca 41580
 agtgtgcacg agcaatagct tgcctttacg ttcatattt agcatgtact actactaact 41640
 30 acgcaaaaat caattcacgg attattaaac tgttaacatc attttagcac gttaacatat 41700
 gtctattca acacaccgtt ttggcacat ttacaaactt gcaaagtigc aatactccct 41760
 tegtacata gcataagaga ttttaggtga atgtacaca tctatccaaa ttcatattac 41820
 35 tagaatgtat caccgcctcc acgccgggag ggagagcgcc gccggtggag aaagggggag 41880
 ggagtggtcg aggggaacca gtagggtgcc ctccccgtcg ccgcctcccc gtggccgcgc 41940
 cggcgagaca ggaggaagag ggggataagg agcggcgccg ccggtgaggg cgcgcgcgcg 42000
 40 ggggggagcg gcgacgccgg tgaggaaggg aaggggagtg gtggctttga gagagatagg 42060
 ggggaggaaa aatgatttta gagttaggtt ttgggtgctt gagttttat atagatcggg 42120
 atcaatcagg accgtccatc agatcggaca actacggctt ctccgcgtt gggccgggtg 42180
 45 ccactccctag gtgtccaca ctattgggcc acatgtacgc tccgcgtgaa ataagttcac 42240
 tttaggtcct ttaagttgcc tctgaattgt tcccaggccg gccgcactat tgggccaccc 42300
 50 cataggccat gtgtacgcic cgcacagaat aatttcgtt tagctccctt aattgtccc 42360
 ctcaaactcc taaaaccagt gcaaatcttt aatttttagt tcacccattg caactcacgg 42420
 gcataattgc tagtgacata taalatgaaa cgaaggatgt agcagactat agaatttaaa 42480

ctgtgctttc attttagagc atcactaact gttattttaga tttttattta aataaatgct 42540
 gaaatgatgt ttttattatg aaaatttagca ataaagctcc caaaatttca aaaaaaaatt 42600
 5 aaaagagatt tatlaatcat ggtaatttta attaaaaatt aaatctaacc atatcatatt 42660
 atttcacggt ccgtgatgag gaaatggcag ctgctatcac ttacggtggt agagaagggg 42720
 10 cattgtttat ttttataact atctcttata actcccatga aactataaaa taaatataat 42780
 cattatcata acattagttt tttttccatt gcaacgcaag ggtaattttt cagtacaata 42840
 aaaaaaataa aagtgggcca ttctgaacgg aaatttctgg ttttttttcc caagagcgcc 42900
 15 gcacacaact gcgcaagaga tcatcgcgga tcacctgct cgtcgccgat ctccacacc 42960
 atccctgcca tctcttccc ctccactggc tctgtctgca cctgtcagct agggcgggca 43020
 tggcgcgccg cgccgcttcc cgcgtctgtg gcgcccttgc ctgggagggc tcatccaag 43080
 20 ggcgaggggg ccgcgcgggg ggcagtggcg gtggcgcgga ggacgcacgc cactgttctg 43140
 acgaattgct ccgtcgtggc ataccagatg tcttctccta caatattctt ctcaacgggc 43200
 tgtgtgatga gaacagaagc caagaagctc tcatgttact gcacataatg gcgatgatg 43260
 25 gaggtagctg cccactgat gtgggtctgt acagcacctg catcaatggc ttcttcaagg 43320
 aggggggatct ggacaaaatg ctgaccaga ggatttgcgc aaatgttgt accitacaact 43380
 ctattattgc tgcgtatgc aaggctcaaa ctgtggacaa ggccatggag gtacttacca 43440
 30 ccatggttaa gactgggtg atgcttgatt gcatgacata taatagtatt gtgcatgggt 43500
 tttgtcttcc agggcagccg aaagaggcta ttgtatttct caaaaagatg cgcagtatg 43560
 35 gtgtcgaacc agatgttgt acttataact cgctcatgga ttatctttgc aagaacggaa 43620
 gatgcacgga agcaagaaag attttgatt ctatgaccaa gaggggccta aagcctgata 43680
 ttactacca ttgtacctg ctccaggggt atgtaccaa aggagccctt gttagatgc 43740
 40 atggctctct ggatttgatg gtacgaaacg gtatccacc taatcattat gtttcagca 43800
 ttctagtatg tgcatacgt aaacaagaga aagtagaaga ggcaatgctt gtattcagca 43860
 aaatgaggca gcaaggattg aatccgaatg cagtaccia tggacagtt atagatglac 43920
 45 ttgcaagtc aggtagagta gaagatgcta tgccttatt ttgagcagat atcatgaag 43980
 gactaagacc tgacagcatt gtttataact ccctaattca tagtctctgt atctttgaca 44040
 50 aatgggagaa ggcgaagag ttatttcttg aaatgttga tggaggcatc tgccttagca 44100
 ctattttctt taattcaata attgacagtc atgcaaaga agggaggggt atagaatctg 44160
 gaaaactctt tgacttgatg gtacgaattg gtgtgaagcc cgatatcatt acccttggca 44220

ggtaagaatgg atgaagcaat gaagtiacii tctggcatgg tctcagttgg gtigaaacct 44280
 aatactgtta cttatagcac ttigattaat ggctactgca aaattagtag gatggaagac 44340
 5 gcgttagtic tttttaagga gatggagagc agtggigtta gtccatgatat tattacgtat 44400
 aacataatic tgcaaggiti atticaaacc agaagaacig ctgctgcaaa agaactctat 44460
 10 gtcaggattta ccgaaagtgg aatgcagati gaactitgti agatttaati ggataattaa 44520
 tccattttaa tcaattlaaat caaataaati ccaaggctca ttatgctagg aattcatgtg 44580
 aattcattct tctatgggat atcaatggga tgaagagiti tgagaattaa tccatttgat 44640
 15 taaggaattg glaacttata tcaattaatc ctatgtatg gatggttgat ggltgtgtag 44700
 tggaggatgg ttcatggcta gttagagaca attagtgtct ctatccctct tccatattca 44760
 ttggtaactt acatcaattta ctcttaattg attgttgggt gatggttgtg tagtggagga 44820
 20 tggttcatgg ctatgtatg acaattagti gctccattcc tcttccctatt ccatgactct 44880
 tactcttcat ctccattcc tcttataaaa tgagaatgga ttgatctcc cgcgagaaga 44940
 agaagacaca ctttcatcca ttttcaaaag ctgttgcctg tacggtaatc ccatcccgac 45000
 25 gagtgtgtgc acacgcgttg ggagagttag cctccgaac cacgcgttgc tgcgacgtti 45060
 gcacagacgg gcgggcgatc aggtttttgg ggagcgcaag gcgcgactac tcactgttcg 45120
 tcaacatcta cttaacttc accaacaigt cgaacactgg agacaaggag aaggagactc 45180
 ccgtcaacac caacggagge aatactgctt caaactccag cggaggacca ttcttggggt 45240
 ataaccttat tacattatit caattagaag tttaactgtt aatgttcatc gcaatgtcaa 45300
 35 catgtgtica ttatgtatg gttagatgct attcaacgtt aagcatgctc atgttgattta 45360
 cattcaccac tatcatgga tcaaatctta ttglaaatat catgtttatt atcttgittat 45420
 tttagattaa aatatgccga attatgacca aatttccaac aaacttagca catacaacat 45480
 40 aatccittcat ggactttgca aaaacaaact cacgatgat gcacttcgaa lgtttcagaa 45540
 cctatgtttg atggatttga agcttgaggc taggacttcc aacattatga ttgatgcatt 45600
 gcttaaagti ggcagaaatg atgaagccaa ggatttgtti gtgtcttctt cgtctaacgg 45660
 45 tttagtgccg aattattgga cgtacagatt gatggctgaa aatattatag gacaggggtt 45720
 gctagaagaa ttggatcaac tcttcttctt aatggaggac aatggctgta ctgttgactc 45780
 tggcatgcta aatttcattg ttagggaact gttagcagaga ggtgagataa ccagggttgg 45840
 50 cacttacctt tccatgattg atgagaagca ctttccctc gaagcatcca ctgcttctt 45900
 gtttatagat ctttgtctg ggggaaaata tcaagaatat catatatttc tccctgaaaa 45960

atacaagtc tttatagaat ctltgagctg ctgaagcatt ttgcagcttt gaaattctgt 46020
 gttggaatic tttctccta cagtcggatt agaggaggga tcttctctgt atgtgtaaat 46080
 5 agcgagggat gtaatgcacc tctccgaatt attltgacig tggtlcctgg actgtaaaca 46140
 agclattatc tcttggtgtt gatgccagaa aaaacacaaa agtttgtctt tatctctact 46200
 10 aacggatcat aaaggggttt gtaactggag ttcaaacctt aagglatcta ggcagtaggt 46260
 alalattgat cctacatctt atgaltctaa gatgalaacc tctcattat cctctgctga 46320
 aacittagct tgaaccgtca tctacaccac aatttgagcc ccttagcaca gagcacaacg 46380
 15 agcaatagct tgccttacg ttcattattt agcatgcact actactaact acccaataat 46440
 caatatactg gttatlaaac tgtttgtaca gtttaataat gtcaatttat cacgttaaca 46500
 taigtittcat tcaacaccac accggttttg gcacagtigc aaacttgcaa taacattttt 46560
 20 actacttctc cgccccataa tataacaatc tcttccata ctatatgtct atattacggg 46620
 acggatgaag tacttcttct ctcccaaaat ataagaatct agtcctagat tagatattat 46680
 ttggattcac gaatttgatt aggctatcta gatttgtagt cgtatgtaat gictaattcg 46740
 25 gtaatagggt attacctctt tggatggagg gagtgtttt tatttctgtac tccctctgtt 46800
 tcatattata agttgttttg acttttttct tagtcaaatt ttattgagtt tgactaaatt 46860
 tatagaaaaa aaattagcaa catttaagca ccacattagt ttcaataaat gtagcatgga 46920
 30 atatattttt ataatagttt tgttttttta ttaaaatgct actatatttt tctataaatg 46980
 tagccaaatt taaagaagtt tgattacgaa aaaaaatcaa aatgacatai aatatgaaac 47040
 35 tgaggatgta gcagactata gcaaatttaa actatgcttt tatttttagag catcaccaaa 47100
 agagatagcc taaactttat cttaactaat taaaataltc ataattttcc ttctgtcaca 47160
 ttaaattttc gtccglaaat ccgattgaaa tccaactaga caatccaaaa aatagagaaa 47220
 40 aagaacagaa aaaataataa aaagcacaca aatcttatct caatcccgcg ggaagctgcc 47280
 gatgccgccg aatccgctcg agcgccgccg ccgcccctca cggggaacga tctcgctgct 47340
 atcgcacgtg glatgggagg gcgcccgcgc cgcgtcttgg gagataggat atggagagag 47400
 45 aaggaaatgt gagggagggg laggtttttc cccattcgta tcttcagcga cacggaggcg 47460
 atccaagctg tccatcagat cagacggctc agaacgcctc catcttcagg ccgcgcattgc 47520
 ttgatgggcc gagggaaggc cggagggtcg aacaaacgta gtcagaggag gatttggagg 47580
 50 aggtaaagta gaatttattt gcgggctgag atagttaaag gactgaaaat ggcccataga 47640
 gaaattggga attttattta aataaagtgt gaaaagggtt ttatatattc aaaattagaa 47700

5 attaagctcc gaaaatttta aaaaataatc aaagagcatt attaatcatg attaatftaa 47760
 taaaaattaa atccaacat atcatattat ttacaggcgc gcagtaggaa aatgcgcagc 47820
 10 tglitgctct tacggiggga gagaaggac atlgtttatt ttcagaacta tcttttataa 47880
 ctcccatgga actttaaaat aaatataatc attattatag cattagtiti ttctgtctt 47940
 15 ttttttcccc aagagcgccg cgcagaagag atcgatcgcg atctccctgc cccgacgtcg 48000
 ccggccgac tctcatctc tccacgccc gtctgtcgcc gatctctac accatccctg 48060
 ccctctctc ctccccctc cctctatct ccactggigc cggccacctc lccgtataag 48120
 20 acaaactgcg ttgcggcggt gggttccgcc ggcgctgctg ctgcacctgt cagctagggc 48180
 gggcatggcg cggcgcgccg ctccccgcgc tgttggcgcc ctctgtctcg acggctcgat 48240
 ccaagggcga ggaggccgcg cggggggcag tggcgccgag gacgcacgcc acgtgttcga 48300
 25 cgaatgtct cggcgiggca gggcgccctc gatctacggc ttgaaccgcg cctctcgcca 48360
 cgctcgcgct cacagccccg cggccgccc gtcccgtac aaccgcatgg cccgagctgg 48420
 cgccgacgag gtaactcccg actgtgcac ctacggcatt ctcatcggti gctgtgtccg 48480
 30 cgcgggccgc ttggacctcg gtctcgcgcc ctltggcaat gtcaitaaga agggatttag 48540
 agtggaaacc atcacctca ctctctgtct caagggccctc tgtccgaca agaggacgag 48600
 35 cgacgcaatg gacatagtc tccgcagaat gaccgagctc ggttgcatac caaatgtctt 48660
 ctcttacaat aatctctca acggcgctgt tgaatgagaac agaagccaag aagctctcga 48720
 gtgtgtcac atgatggctg atgatcgagg aggaggtagc ccacctgatg tgggtgtcga 48780
 40 taccactgtc atcaatggct tcttcaaaga gggggatca gacaaagctt acagtacata 48840
 ccatgaaatg ctggaccggg ggattttacc tgaatgtgtg acctacagct ctattatgc 48900
 tgcgttaatg aagggtcaag ctatggacaa gccatggagg tacttaccac gatgttaag 48960
 45 aatgggtca tgcctgatg catgacatat aatagttatt tcttgaaatg ttggatcgag 49020
 gcatttgtct ggacactatt ttctttaatt caataatga cagtcattgc aaagaaggga 49080
 gggttataga atctgaaaaa ctcttgacc tgaatgtacg tatgtgtgtg aagcctgata 49140
 50 tcattacata cagtacactc atcgatggat attgcttggc aggttaagat galgaagcaa 49200
 tgaagtact tcttggaatg gtctcagttg ggttgaaacc taatactgtt acttatagca 49260
 55 ctltgattaa tggctactgc aaaattagta ggaatggaaga cgcgttagti ctitttaagg 49320
 agatggagag cagtgtgtgt agtccgata ttattacgta taacataati ctgcaagggt 49380
 tatttcaaac cagaagaact gctgtctgcaa aagaactcta tgtcaggat accgaaagt 49440

gaacgcagat lgaacttagc acafacaca taatccitca tggactitgc aaaaacaaac 49500
 tcactgatga tgcacttcag atglttcaga acctatgltt gatggatttg aagcttgagg 49560
 5 ctaggacitit caacattatg atgatgcat tgcitaaagt tggcagaaat gatgaagcca 49620
 aggatitgtt tgtitcttcc tgcitaaag gtttagtgcc gaattatitg acgtacaggt 49680
 10 tgaatggcga aaataatata ggacaggggt tgcitagaaga attgatcaa ctttttcttt 49740
 caatggagga caatggctgt actgttgact ctggcatgct aaatttcatt gtiagggaac 49800
 tgtitcagag aggtgagata accagggctg gcacttacct tccatgatt gatgagaagc 49860
 15 acttttccct cgaagcatcc actgttccct tgtttataga tcttttctct gggggaaaat 49920
 atcaagaata ttataggttt cccccgaaa aatacaagtc ctttatagaa tctttgagct 49980
 gctgaagcat ttgcagctt tgaattctg tgttggaaat cttttctctt acagtcctat 50040
 20 tagaggaggg atcttctctg tatgtttaa tagcgaggta tgaatgccac ctctccgaat 50100
 tatitttact gtgttctcta gactgtaaac aagcaattat gttatgctgt tgaatccaga 50160
 aaaaacataa aagtttgtcg ttaictctac taacggatca taaagggatt tgtactgga 50220
 25 gtittcaaact taatgtgtct aggcagtaat ttgacatta gatccaaaac aatttatagg 50280
 gtittcatiaa atttcatcta tgtgtactgt ttaggtgttg aatagtttga ctgtttttt 50340
 aactgaacaa aagatagtc tgaagcttg ttcittacca aatgcagtac tgaatcatc 50400
 30 aataatattt ttatggaaca agattggatt gtaagaaatg gtttccgac tgaattatct 50460
 atctcaacgt attattatgc acatgtacta atcatgaaat atctgatgga atgatgttct 50520
 35 tatitaccig tgtgaggcag caaggagtga gatggataac accacatact cctctatcc 50580
 cagaatataa gaagttitag agtggacac gattatlaag aaagtaggta gaagtgagta 50640
 gtggaggggt gtgatgcat gagtagtga ggtaggtggg aaaagtgaat ggtggaggggt 50700
 40 tgtgattgggt tgggaagaga atgttggtag agaagtgtt atattitggg gattacatta 50760
 ttattctaga acaatactgt tgtgtcaag aagcgttcca aagaatgttc acaacctgtg 50820
 ctcatgggt tttagactta atcttggac attcagtaic atgatctgc tcatctttaa 50880
 45 acatggaata aaggatgaca gcatgatttc ttgtctcta taatctttg gctaccacaca 50940
 gataatagct gtaaatctat actacitaa aaggagtagt ggtgggtgtg agtggtgaat 51000
 50 ctgccaccac cccaccacca actctcaaaa ttctgacatg tgggatcac gtcaatccct 51060
 tctccaagac atgtgggac actgtcaatc cttcttcaa accaatgtga tgaagaaca 51120
 gtggaaatca cggacagacc atggagctct caaccataat catcttgcg agttaataac 51180

aaatggagcg taaacttggc aagcaaaaaa ctcaaattaa ttctaaaatt aagctctagg 51240
 attcaaaaata gatttctctt ctgcatltgt cigtatgat ttttaattcc gtaacaacgc 51300
 5 aaatgcatit tgctagtcit ataaagaagg gttaatgcaa atattctgat taaatgattg 51360
 tatctatgaa gtttgaatgc tagtggagc tcttttgacc atgttttgtt gtgcgagcat 51420
 10 ttaagagagt gaagagaatg ctcttttggg gctgttctgg tatggaagga tccacagata 51480
 aaattcaggt tctactgcit ctctgttgtt aattttcatg aagctgcagt gaataccttg 51540
 ttgaccacit gatctgttgc ttggaaggag aatatagtag tggccaaggt tggtagcgtt 51600
 15 gatggtagca tgtgatcccc cagatcttca gtgaccaga gaggaggga cggcgcgttg 51660
 tgagctacaa ggcatactca gtggaggga agatcaaggc ctcccgctcg taggggactc 51720
 cgtctcatca aggccaactg ctccgaactg atcaatttct ggtacggatc acttctcctt 51780
 20 tctttttttt ttccaccita agcacitctt tgattcttct cgtctacctc cttaatttcc 51840
 ttccaalata ttgtggcact tgaatcatggc ggagaccac ctccaggtt gaatggattt 51900
 tgtcaaagaa ctaaatttat tccattagct tattttctga ttacatggaa gacattcttt 51960
 25 tctggaataa atacagaact aaatccgtt tctgaataa aagtgttag tgtgtggcat 52020
 ggtgcatttc cgcgcttcta aattttataa aacctgttca ttcaattga acctgcatcc 52080
 30 aatccaatat tttaggtgca gacagggtct tgcggtcagg ttaaagaagt tggcaaaaat 52140
 gcttctgaag aaaggttaat tgttgtttca tctcaggagg taataatcag atgattattc 52200
 caattggcat tgccttgcca tttttatcac gagtctttac aattttatat cctcctacat 52260
 35 attctttcca gattccagat gatccaggtt ctccaacaat tgaggcgctt attttgctcc 52320
 ataglaaagc aagtacacit gctgagaacc accagttgac aacacggctt gtgtaccat 52380
 caaacaaggt tggttgtatt ctgggggaag gtaggaaggt aattactgaa atgagaagac 52440
 40 ggactggggc tgaatccga gtctactcaa aagcagataa acctaaagtac cgtctttttg 52500
 atgaggagct tgtgcaggta atttatttgg ccatacctac accagagatc catatattac 52560
 ttttataact gcagttttta ctgtttaaca ttcatltgt cttttacatt tgttccaagc 52620
 45 ttccaggttg ctgggcttcc agctatlgaa agaggagccc tgacagagat tgcctcgagg 52680
 ctltgaacta ggacactcag agatggaagt tcttccaata atccgacacc ttttggccct 52740
 50 gttagtggtc ctccgttga tatcttgctt aacaaggaat tcatgtata tggacgatct 52800
 gctaatagtc ccccatatgg agggcctgct aatgatccac catatggaag acctgccatt 52860
 gatccaccat atggaagacc aatatccaca atatggaaga ctgccaatg atccaccata 52920

tagaagacct gtcaatgata catcatattg aggggtggac aatgatgggc ctctgatca 52980
 5 ggccccgggtcc tgaggggggt cgaatggggc gatcgctccg ggccccgat tcccagggcc 53040
 cccacclatc lgtgcaacga gtagtagcga tcttccagcg cgcaacgtga ggcgatgttt 53100
 ctccgtgatt tcgccggcct gcaactgcga gatcgcgagt ataacgatca gccgatcgat 53160
 10 ctcatctgcc gactgccatg ctgatgccac acgcaagcgc agcatafcag ccttatcttg 53220
 gttgatcggc atgcgggacg agcacatctg ttgtcgatc aactgcigac tgcctatata 53280
 gtgcctgggtgc tgaatcgatc gattgtctgc gcggaagtga agaacaacca cggcacigct 53340
 15 gccctgcggg ctctagccgc catcagtaag tacgctatc tgcctatcta gatctagatc 53400
 gagattacat agtggaaata tctgtttata acaaaattac aaggtatcaa ttgataattt 53460
 aaggittataa ccgtacaaac ttcatgtatt tgcctgtttc acattggta gatttgtttc 53520
 20 aactaatttg gtacttctgt agccttgtaa ttacgaatc tagtattaat attttcttaa 53580
 gtattagcct gttccttgat attatgctgt tgagaaagta tgcaatagat aacaaaaaca 53640
 25 agtaggtgtg ttgaggatgc tcaagagtaa tacagccact tcaataattc tgatattatc 53700
 aggacatcat caataattct gcgcctacaa atcttcaaag aaaattttaa tataatgcgt 53760
 atgatTTTTT aaatacgaat atgattgct atttaaagat atttatatta tatggtaatt 53820
 30 attatttgaa ggtttataat aaaggccctc gtttttagtt tcacgctggg ctttcagaat 53880
 ctccaggaccg gccctgctca tgatc 53905

35 <210> 29
 <211> 24
 <212> DNA
 40 <213> artificial sequence
 <220>
 <223> Oligonucleotide primer for amplification
 45 <400> 29
 atcaggagcc ttcaaattgg gaac 24

50 <210> 30
 <211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 30

ctcgcaaatt gcttaatttt gacc 24

<210> 31

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 31

tgaaggagtt atgggtgcgt gacg 24

<210> 32

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 32

ttgccgagca cacttgccat gtgc 24

<210> 33

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 33

gcgacgcaat ggacatagtg ctcc 24

<210> 34

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 34

ttaccigcca agcaatatcc atcg 24

<210> 35

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 35

aaggcatacl cagtggaggg caag 24

<210> 36

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 36

ttaacctgac cgcaagcacc tgtc 24

<210> 37

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 37

tggatggact atgtggggtc agtc 24

<210> 38

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 38

agtggaagtg gagagagtag ggag 24

<210> 39

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 39

ccctccaaca cataaatggt tgag 24

5 <210> 40

<211> 24

<212> DNA

10 <213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

15 <400> 40

tttctgccag gaaactgtta gatg 24

20 <210> 41

<211> 24

25 <212> DNA

<213> artificial sequence

<220>

30 <223> Oligonucleotide primer for amplification

<400> 41

gcgatcttat acgcatacta tgcg 24

35 <210> 42

40 <211> 24

<212> DNA

<213> artificial sequence

45 <220>

<223> Oligonucleotide primer for amplification

<400> 42

50 aaagtccttg ttccttcacc aagg 24

55 <210> 43

<211> 26

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 43

gaggatttat caaaacagga tggacg 26

<210> 44

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 44

tgggcggcag cagtggagga taga 24

<210> 45

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 45

aagaaggag ggttatagaa tctg 24

<210> 46

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 46

atatcaggac taacaccact gctc 24

<210> 47

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 47

acgagtagta gcgaiccttc agcg 24

<210> 48

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 48

cagcgtgaaa ctaaaaacgg aggc 24

<210> 49

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 49

atccacatc atcataatcc gacc 24

<210> 50

<211> 25

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 50

agcttctccc ttggatacgg tggcg 25

<210> 51

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 51

atttgttggg tagttgcggc tgag 24

<210> 52

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 52

gccc aaactc aaaaggagag aacc 24

5

<210> 53

<211> 24

10

<212> DNA

<213> artificial sequence

<220>

15

<223> Oligonucleotide primer for amplification

<400> 53

cctcaagtct cccctaaagc cact 24

20

<210> 54

<211> 24

25

<212> DNA

<213> artificial sequence

<220>

30

<223> Oligonucleotide primer for amplification

<400> 54

gctctactgc tgataaaccg tgag 24

35

<210> 55

<211> 24

40

<212> DNA

<213> artificial sequence

45

<220>

<223> Oligonucleotide primer for amplification

50

<400> 55

tggatggact atgtggggtc agtc 24

55

<210> 56

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 56

agtggaagtg gagagagtag ggag 24

<210> 57

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 57

tacgacgcca ttccactcca ttgc 24

<210> 58

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 58

catttctcta tgggcgttgc tctg 24

<210> 59

<211> 26

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 59

accgtaggt atggcacctt caacac 26

<210> 60

<211> 26

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 60

ccaaggaacg aagttcaaat gtatgg 26

<210> 61

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 61

tgatgtgttt gggcatccct ttcg 24

<210> 62

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 62

gagatagggg acgacagaca cgac 24

<210> 63

<211> 26

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 63

tcctatggct gtttagaaac tgcaca 26

<210> 64

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 64

caagttcaaa cataactggc gttg 24

<210> 65

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 65

cactgttcctg taagtgctgct glgc 24

<210> 66

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 66

caagcgtgtg ataaaatgtg acgc 24

<210> 67

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 67

tgccctactgc cattactatg tgac 24

<210> 68

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Oligonucleotide primer for amplification

<400> 68

acatactacc gtaaatggtc tctg 24

Claims

1. A method for restoring rice fertility comprising introducing a nucleic acid into rice, wherein the nucleic acid has the base sequence of SEQ ID NO.27, or has a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27, and which functions to restore fertility.
2. A method for restoring rice fertility comprising introducing a nucleic acid into rice, wherein the nucleic acid has the base sequence of bases 38538-54123 of SEQ ID NO.27, or has a base sequence which is identical to at least 70% of the base sequence of bases 38538-54123 of SEQ ID NO.27, and which functions to restore fertility.
3. A method for restoring rice fertility comprising introducing a nucleic acid into rice, wherein the nucleic acid has the base sequence of bases 42132-48883 of SEQ ID NO.27, or has a base sequence which is identical to at least 70% of the base sequence of bases 42132-48883 of SEQ ID NO.27, and which functions to restore fertility.
4. The method of any one of Claims 1-3, wherein the nucleic acid having a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27 or of the base sequence of bases 38538-54123 of SEQ ID NO.27, and which meets at least one of the following requirements 1) and 2):
 - 1) a base corresponding to the base 45461 of SEQ ID NO.27 is A; and
 - 2) a base corresponding to the base 49609 of SEQ ID NO.27 is A.
5. The method of any one of Claims 1-4, wherein the nucleic acid comprising the rice restorer gene locus introduced into rice, does not comprise any constitutional gene other than the rice restorer gene.
6. A method for discerning whether or not a subject rice individual or a seed thereof has the rice restorer gene (the Rf-1 gene) or not, wherein the method utilizing a fact that a sequence determining the presence of the function of the Rf-1 gene positions between the polymorphism detection marker loci P4497 MboI and B56691 Xba I on rice chromosome 10.
7. The method of Claim 6 wherein the Rf-1 gene exists in a nucleic acid having the base sequence of SEQ ID NO.27, or a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27.
8. The method of Claim 6 wherein the Rf-1 gene exists in a nucleic acid having the base sequence of bases 38538-54123 of SEQ ID NO.27, or a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of bases 38538-54123 of SEQ ID NO.27.
9. The method of claim 6 wherein the Rf-1 gene exists in a nucleic acid having the base sequence of bases 42132-48883 of SEQ ID NO.27, or a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of bases 42132-48883 of SEQ ID NO.27.
10. The method of any one of claims 7-9 wherein the subject rice individual or the seed thereof is determined to have the Rf-1 gene, in the case that the nucleic acid having a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27 or of the base sequence of bases 38538-54123 of SEQ ID NO.27, meets at least one of the following requirements 1) and 2):
 - 1) a base corresponding to the base 45461 of SEQ ID NO.27 is A; and
 - 2) a base corresponding to the base 49609 of SEQ ID NO.27 is A.
11. The method of any one of Claims 7-10, wherein the method comprises:
 - i) preparing a pair of primers based on a base sequence of adjacent regions including any one of the following base;
 - 1) a base corresponding to the base 45461; or
 - 2) a base corresponding to the base 49609; to amplify both the base of the above and adjacent regions thereto;

ii) performing nucleic acid amplification reaction(s) using the genome DNA of the subject rice individual or the seed thereof as a template; and
 iii) discerning the presence of the Rf-1 in the subject rice individual or the seed thereof based on polymorphism found in said nucleic acid amplification product.

5

12. The method of Claim 11 wherein the subject rice individual or the seed thereof is determined to have the Rf-1 gene, in the case that the step iii) meets least one of the following requirements 1) and 2):

10

- 1) a region including a base corresponding to the base 45461 of SEQ ID NO.27 does not have any TaqI recognition sequence; and
- 2) a region including a base corresponding to the base 49609 of SEQ ID NO.27 does not have any BstU recognition sequence.

15

13. The method of Claim 11 or 12 using a pair of primers having base sequences selected from the group consisting of SEQ ID NO:45 and 46, and SEQ ID NO:47 and 48.

20

14. A method for inhibiting the function of the Rf-1 gene to restore fertility by introducing an antisense having at least 100 continuous bases in length, and having a base sequence complementary to a nucleic acid having the base sequence of SEQ ID NO.27, or to a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27, and which functions to restore fertility.

25

15. A method for inhibiting the function of the Rf-1 gene to restore fertility by introducing an antisense having at least 100 continuous bases in length, and being selected from base sequences complementary to a nucleic acid having the base sequence of bases 38538-54123 of SEQ ID NO.27, or to a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of bases 38538-54123 of SEQ ID NO.27, and which functions to restore fertility.

30

16. A nucleic acid having the base sequence of SEQ ID NO.27, or a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of SEQ ID NO.27, and which functions to restore fertility.

35

17. A nucleic acid having the base sequence of bases 38538-54123 of SEQ ID NO.27, or a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of bases 38538-54123 of SEQ ID NO.27, and which functions to restore fertility.

40

18. A nucleic acid having the base sequence of bases 42132-48883 of SEQ ID NO.27, or a nucleic acid having a base sequence which is identical to at least 70% of the base sequence of bases 42132-48883 of SEQ ID NO.27, and which functions to restore fertility.

45

50

55

Fig. 1

RFLP Probe (S12564)



Probe A ■ Probe E ■ Probe F ■ Probe G ■ Probe H ■

WSA 1

WSA 3

WSE 8

WSF 5

WSF 7

WSG 6

WSG 2

(1) Asominori

XSE 1

XSE 7

2618

17261

12478

XSF 4

33750

XSF 20

20328

XSG 22

31694

XSG 16

38538

54123

XSO 8

46538

XSH 18

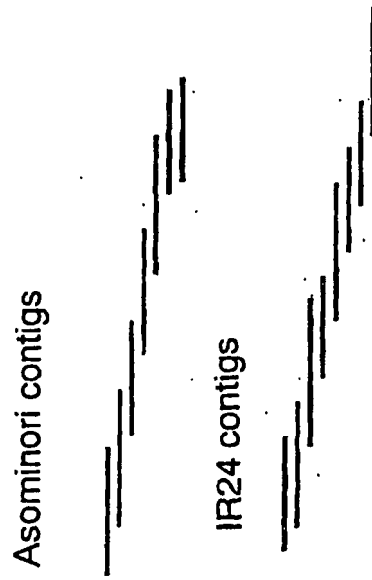
63364

56409

76363

(2) IR24

Fig. 2



AC068923

Fig. 3

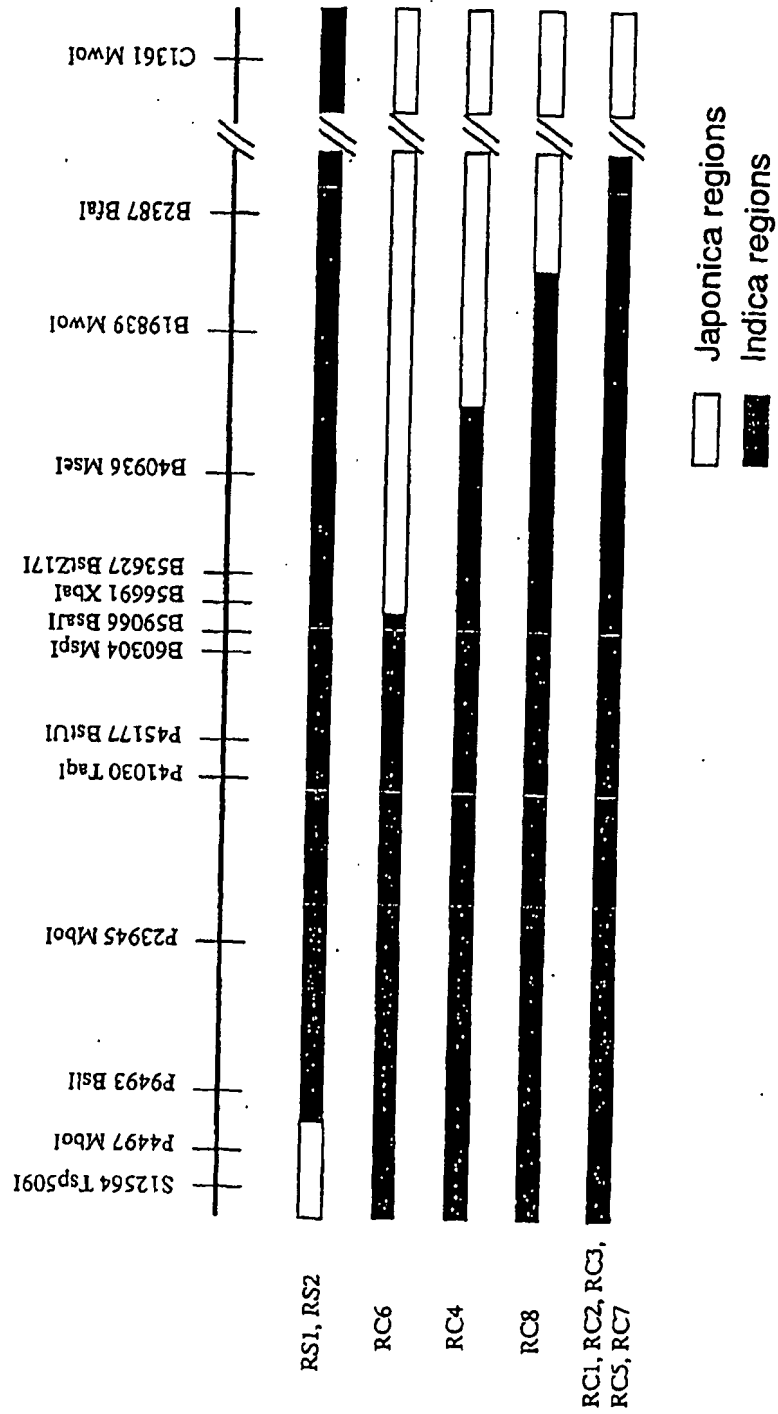


Fig. 4

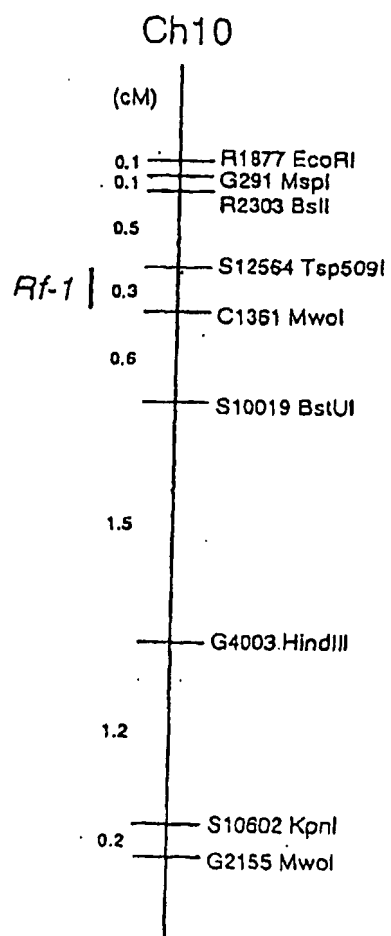


Fig. 5

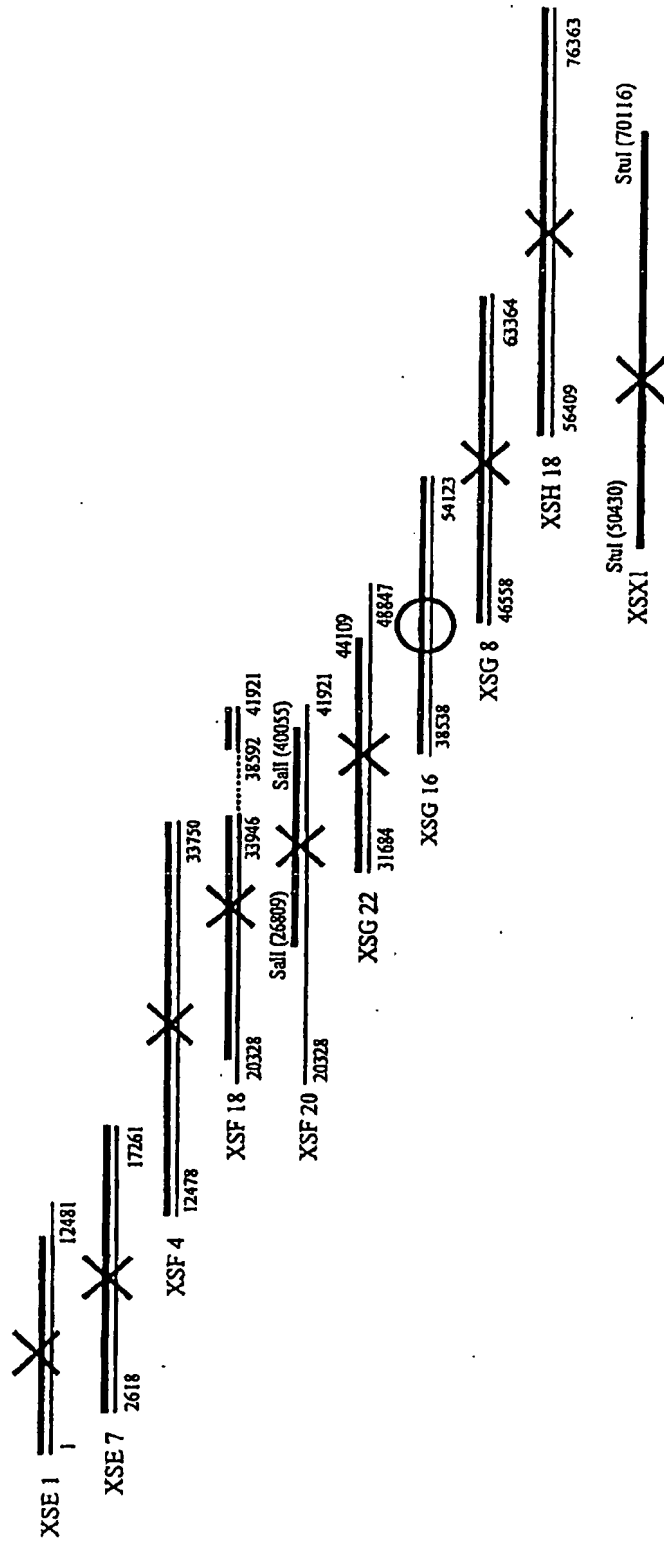
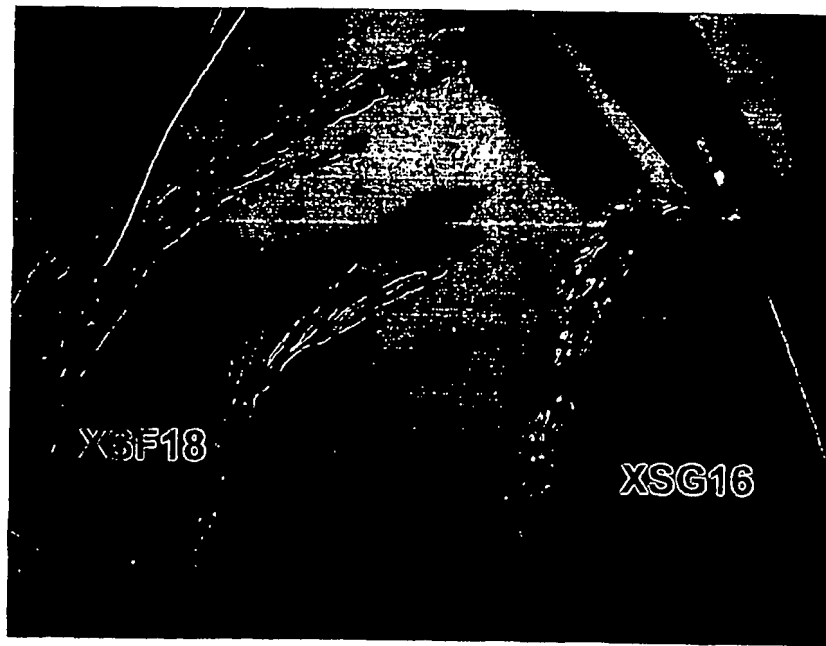


Fig. 6



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP02/09429

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ C12N15/29, C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ C12N15/29, C12Q1/68

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Genbank/EMBL/DBJ/GeneSeq, WPI (DIALOG)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Yoshimichi FUKUDA et al., "RFLP Marker o Mochiita Ine no Nensei Kaifuku Idenshi(Rf-1) ni Kansuru Rensa Bunseki", Breeding Science, 1992, Vol.42, suppl.1, pages 164 to 165	1-18
A	JP 9-313187 A (Mitsui Toatsu Chemicals, Inc.), 09 December, 1997 (09.12.97), (Family: none)	1-18
A	JP 2000-139465 A (Mitsui Chemicals, Ltd.), 23 May, 2000 (23.05.00), (Family: none)	1-18
P,A	JP 2002-345485 A (Japan Tobacco Inc.), 03 December, 2002 (03.12.02), (Family: none)	1-18

☐ Further documents are listed in the continuation of Box C.
 ☐ See patent family annex.

* Special categories of cited documents:	"I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
21 November, 2002 (21.11.02)Date of mailing of the international search report
10 December, 2002 (10.12.02)Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.